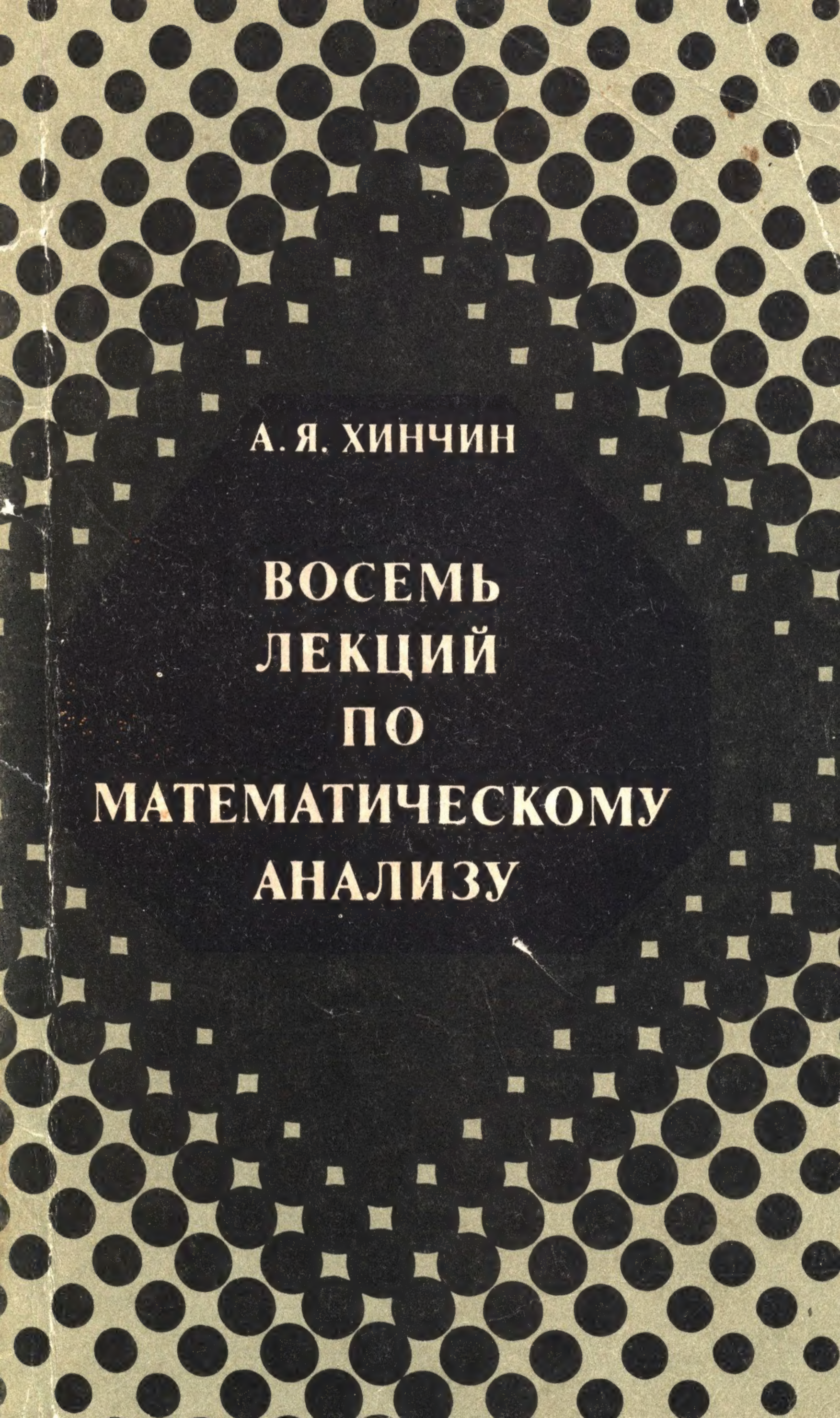

А. Я. ХИНЧИН
ВОСЕМЬ
ЛЕКЦИЙ
ПО
МАТЕМАТИЧЕСКОМУ
АНАЛИЗУ

А. Я. ХИНЧИН

ВОСЕМЬ ЛЕКЦИЙ ПО МАТЕМАТИЧЕСКОМУ АНАЛИЗУ

ИЗДАНИЕ ЧЕТВЕРТОЕ



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
Москва 1977

517.2
Х-47
УДК 517

АННОТАЦИЯ

Книга посвящена изложению ряда принципиальных вопросов математического анализа, которым часто в курсах высшей математики уделяется мало внимания. Книга обращена к читателю, уже обладающему познаниями по математике примерно в объеме курса втуза; она будет полезна инженерам, преподавателям вузов, учителям средних школ и студентам-математикам.

$Х\frac{20203-000}{053(02)-77}$ 24-77

СОДЕРЖАНИЕ

ОТ РЕДАКТОРА

Один из наиболее выдающихся математиков своего времени, Александр Яковлевич Хинчин (1894—1959), известный своими работами по теории вероятностей, теории чисел и теории функций, всю жизнь глубоко интересовался преподаванием математики. Среди книг, принадлежащих его перу, совсем особое место занимают его «Восемь лекций по математическому анализу».

Эта книга написана для инженеров, учителей и экономистов, для всех тех, кто изучал математический анализ и в общем знаком с технической его стороной, с формулами, но хотел бы вникнуть в принципиальные вопросы, осознать смысл основных понятий и идей.

И книга прекрасно выполняет свое назначение. Не стремясь охватить весь материал, входящий в полные учебники, А. Я. Хинчин не жалеет времени для того, чтобы выделить и объяснить наиболее существенные вопросы, обычно представляющие главную трудность для изучающих анализ, выявить самую суть рассматриваемых понятий.

Быть может, специалистам, знакомым с современным изложением анализа, некоторые места в книге покажутся несколько старомодными. Но и эти места, давая яркий отпечаток времени, когда появилась книга, написанные с исключительным мастерством, с неподдельной искренностью и блеском, живым и богатым языком, только подчеркивают достоинства книги.

Впервые книга была издана в 1943 г. Она сыграла значительную роль в распространении математических знаний. Предлагая читателю после почти тридцатилетнего перерыва четвертое издание этой книги, мы уверены, что она принесет пользу всем желающим глубже изучить математический анализ, а лучшие ее страницы доставят удовольствие многим любителям математики.

М. А. Крейнес

ПРЕДИСЛОВИЕ АВТОРА К ТРЕТЬЕМУ ИЗДАНИЮ

Третье издание отличается от первого лишь немногими изменениями. Самое существенное из них состоит в том, что я вычеркнул «принцип индукции» из числа основных лемм, вследствие чего все опиравшиеся на этот принцип доказательства пришлось заменить другими. Я надеюсь, что для большинства читателей я этим облегчил усвоение книги, так как мне представляется, что этот принцип и опирающиеся на него рассуждения предъявляли читателю в отношении логической культуры требования несколько более высокие, чем это вообще принято в настоящей книге.

Из других изменений заслуживают быть отмеченными только новая трактовка формулы Тэйлора и параграфа о функциях с ограниченным изменением.

ПРЕДИСЛОВИЕ АВТОРА К ПЕРВОМУ ИЗДАНИЮ

Курс высшей математики в том или ином объеме читается в значительном большинстве высших учебных заведений. С основами математического анализа знакомятся студенты всех технических и большей части военных, агрономических и экономических вузов, не говоря уже о специальных факультетах университетов и педагогических институтов. Все эти многочисленные, читаемые в разных учебных заведениях курсы весьма различны как по своему объему, так и по той принципиальной, идейной и логической базе, на которой они строятся. В последнем отношении, в сущности, только университетские курсы поднимаются до более или менее солидного научного уровня; все остальные вынуждены довольствоваться упрощенным, далеко отстающим

от современных научных воззрений идейно-логическим фундаментом, а иногда и весьма ограниченной программой.

Между тем мы часто встречаемся с таким положением, когда инженер, учитель, экономист, в свое время изучавший высшую математику по такому упрощенному курсу, начинает ощущать потребность в расширении и, главное, более прочном обосновании своих математических знаний. Возникает ли эта потребность вследствие тех или других конкретных исследований такого специалиста в его собственной научной области, или она встает как неизбежный плод общего расширения его научного и жизненного горизонта,— независимо от этого она, конечно, должна быть удовлетворена. Казалось бы, это сделать просто: взять какой-нибудь полный курс математического анализа, вроде Гурса [1]) или Валле-Пуссена [2]), и штудировать его систематически, опираясь на уже имеющиеся менее солидные и глубокие познания. Однако опыт показывает, что этот представляющийся столь естественным путь почти никогда не приводит к цели и, за редкими исключениями, ведет к разочарованию, а подчас и парализует дальнейшие попытки в намеченном направлении. Дело в том, что, с одной стороны, наш учащийся, как правило, располагает для той цели, которую он себе поставил, лишь весьма ограниченным количеством времени и систематически проработать большой курс в нескольких томах уже по этой причине лишен возможности. С другой же стороны,— и это, пожалуй, всего важнее,— не имея еще твердой научной базы и не будучи математиком по специальности, он сам, без руководства, не может выделить в изучаемом принципиальных моментов, вынужден уделять поэтому полную долю внимания не имеющим существенного значения мелочам и кончает тем, что теряется в этих мелочах, переставая, так сказать, видеть лес за деревьями.

---

[1]) Э. Гурса, Курс математического анализа. тт. 1—3, ГТТИ, 1933.

[2]) Ш. Ж. де ла Валле-Пуссен, Курс анализа бесконечно малых, тт. 1, 2, ГТТИ, 1933.

А между тем, для того чтобы дать потребностям и запросам такого учащегося полное удовлетворение, нужно весьма немногое. Несколько лет назад я имел случай прочитать специальный, посвященный этой цели курс, состоявший всего из 12 двухчасовых лекций, слушателям курсов по повышению математической квалификации инженеров при Московском государственном университете. Должен признаться, что вначале мне эта задача казалась почти безнадежной; между тем я имею основания полагать, что курс мой удовлетворил запросы слушателей, несмотря на свою краткость. Секрет этого успеха заключался в том, что мне удалось найти правильный ключ к решению стоявшей передо мной педагогической задачи: я с самого начала отказался от мысли излагать хотя бы одну главу своего предмета в полной подробности; вместо этого я ограничивался возможно выпуклым, конкретным и впечатляющим развитием принципиальных моментов, я говорил больше о целях и тенденциях, о проблемах и методах, о связях основных понятий и идей анализа между собой и с приложениями, чем об отдельных теоремах и их доказательствах. Я не боялся в целом ряде случаев для ознакомления с не имеющими принципиального значения деталями доказательств (а иногда и с целыми цепями теорем и доказательств) отсылать моих слушателей к учебнику. Но зато для освещения какого-либо понятия, метода или идеи, имеющих ведущее, принципиальное значение, я уже позволял себе не жалеть времени, стараясь всеми средствами, путем самых разнородных описаний, наглядных образов и т. п. возможно ярче и эффективнее внедрить эти основоположные моменты в сознание моих слушателей. Я с основанием рассчитывал на то, что после такой подготовки каждый из них, ощутив желание или потребность в более глубоком изучении той или другой главы анализа, сможет уже самостоятельно, без посторонней помощи, во-первых, найти нужный ему материал и, во-вторых, подойти к его изучению, так сказать, по-хозяйски, умея отделить главное от второстепенного и малосущественного.

Все многократные беседы мои с отдельными слушателями, с группой слушателей и со всей аудиторией

в целом неизменно убеждали меня, что избранный мною путь был единственно правильным. Пользуюсь случаем отметить, что очень большая аудитория, наводнившая все читавшиеся на этих курсах лекции и в своем большинстве не отсеявшаяся до самого их окончания, наилучшим образом свидетельствует о том, как широко распространена среди наших инженеров потребность в повышении уровня их математических знаний.

Предлагаемая книга ставит себе ту же цель, что и мой только что упомянутый лекционный курс, и стремится осуществить ее теми же самыми средствами. Читатель, следовательно, с самого начала должен быть предупрежден, что здесь он не найдет сколько-нибудь полного изложения университетского курса анализа или хотя бы отдельных избранных глав этого курса. Автор ставит своей задачей только дать общий, но зато возможно доступный и запоминающийся очерк основных идей, понятий и методов математического анализа — такой очерк, который легко бы читался и усваивался всяким, кто знаком хотя бы с самым упрощенным изложением этого предмета, и который, после такого усвоения, уже позволил бы читателю сознательно и самостоятельно изучить во всех деталях любой отдел и любую главу этого предмета.

Я беру на себя, однако, смелость думать, что чтение этой книги сможет принести существенную пользу и многим студентам математических факультетов наших университетов. Дело в том, что ни учебник, ни лектор, будучи, естественно, стеснены рамками времени и программы, не могут уделить достаточно внимания дискуссии принципиальных вопросов и в основном должны ограничивать себя изложением всех деталей конкретного материала. А между тем всякий знает, как полезно иногда оторваться от деревьев и поглядеть на лес; я хочу надеяться, что предлагаемая книга поможет в этом кое-кому из будущих математиков, впервые приступающих к изучению математического анализа.

ЛЕКЦИЯ I

КОНТИНУУМ

Почему математический анализ должен начинаться с изучения континуума?— Почему континуум не может быть изучен без построения полной теории вещественных чисел?— Построение иррациональных чисел.— Теория континуума.— Основные леммы.

Почему математический анализ должен начинаться с изучения континуума? «Переменная величина y называется функцией переменной величины x, если каждому значению величины x соответствует единственное определенное значение величины y». Эту фразу можно уподобить воротам, открывающим вход в область высшей математики; с помощью этой фразы мы определяем важнейшее, основоположное понятие математического анализа — понятие *функциональной зависимости*, в котором, как в зародыше, уже заложена вся идея овладения явлениями природы и процессами техники с помощью математического аппарата. Вот почему мы должны со всей беспощадностью требовать от этого определения полной, безукоризненной ясности; ни одно слово в нем не должно вызывать и тени сомнения; малейшая двусмысленность здесь грозит сделать все величественное здание, которое строит наука на базе этого основного понятия, несовершенным, требующим капитальной перестройки.

А между тем та краткая формулировка, с которой мы начали, при ближайшем рассмотрении оказывается во многих пунктах недоговоренной и допускающей различные толкования. Мы здесь остановимся только на одном таком неотчетливом пункте, ибо как раз попытка выяснить его содержание до конца и

приведет нас вплотную к предмету нашей сегодняшней лекции.

Наше определение содержит слова «каждому значению величины x»; для того чтобы у нас не оставалось никакой неясности, мы должны, следовательно, неукоснительно требовать, чтобы нам был разъяснен смысл термина «значение переменной величины»; но этого мало; в нашем определении говорится о «каждом значении»; отсюда следует, что для полного понимания этого определения недостаточно иметь представление о некоторых отдельных значениях величины x; мы должны, сверх того, совершенно отчетливо представлять себе *всю* совокупность, *весь* «запас» этих значений, любому из которых должно соответствовать определенное значение величины y. Мы должны знать то, что в анализе называют «областью определения» данной функции.

Что представляют собой отдельные значения переменной величины? Мы знаем, что это — числа. Если так, то совокупность всех этих значений есть некоторая совокупность чисел — некоторое, как говорят, *числовое множество*. Какое же это множество, какие числа оно содержит? Мы с самого начала исключим из рассмотрения область мнимых чисел, предполагая, что все значения величин x и y — числа вещественные. Но все ли вещественные числа могут быть значениями величины x? А если нет, то какие могут и какие нет?

Обо всем этом в нашем определении ничего не сказано, и это вполне понятно, потому что на этот вопрос нельзя одинаковым образом ответить для всех функций (а в сущности, даже и для одной функции в разных задачах). Область определения функции зависит как от природы этой функции, так и от тех специальных задач, для решения которых нам эта функция нужна в данном исследовании; ведь очень легко понять, что одну и ту же функцию в разных задачах приходится исследовать при различных значениях независимой переменной. Всему этому мы знаем много примеров. Так, функцию $y = x!$ (по крайней мере в пределах элементарной математики) имеет смысл рассматривать лишь для целых положительных x; функцию $y = \lg x$ — только для $x > 0$ и т. д. Можно легко придумать и такие примеры,

где естественной областью определения функции будет служить числовое множество гораздо более сложного строения.

Однако если мы спросим себя, каково же множество значений величины x, с которым фактически чаще всего приходится иметь дело как в самом математическом анализе, так и в его приложениях, то мы должны будем сказать, что в подавляющем большинстве случаев такой областью определения функции служит «отрезок» (закрытый[1]) или открытый), т. е. совокупность *всех* вещественных чисел, заключенных между двумя данными числами (включая или исключая последние). Иногда этот отрезок заменяется полупрямой (например, $x>0$); это значит, что множество значений величины x есть совокупность *всех* вещественных чисел, которые больше (или меньше) некоторого числа (иногда требование $>$ или $<$ заменяется требованием $\geqslant$ или $\leqslant$); наконец, бывают случаи, когда отрезок обращается в полную прямую, т. е. когда значениями величины x могут быть *все* вещественные числа. В этом случае говорят, что областью определения функции служит вся вещественная ось или «числовая прямая».

Так или иначе, мы видим, что для функций математического анализа той областью, тем полем, на котором они развивают и развертывают свои индивидуальные черты, той средой, где они вырастают в мощные математические орудия естествознания и техники, является совокупность всех вещественных чисел. Эту совокупность в математике называют *континуумом* (точнее — *линейным континуумом*). И совершенно так же, как мы должны тщательно исследовать почву, прежде чем дать ей выращивать нужные нам растения, мы в высшей математике, желая быть рачительными и научно, а не наудачу действующими хозяевами, прежде чем строить на базе понятия функциональной зависимости все здание этой науки, должны тщательнейшим образом изучить ту среду, в которой живет и развивается это понятие. Вот почему во всех серьезных, научно построенных

---

[1]) Обычно отрезок называется не закрытым, а *замкнутым*.— *Прим. ред.*

курсах математического анализа континуум является первым предметом изучения; и только после того, как его сущность и природа в достаточной степени усвоены, можно перейти к основательному изучению понятия функциональной зависимости. А строение континуума оказывается не таким простым, как можно было бы предполагать на первый взгляд; мир вещественных чисел развертывает себя перед нашими глазами как сложная, богатая разнообразнейшими деталями конструкция, всестороннее изучение которой наука не может считать законченным и до настоящего времени.

Почему континуум не может быть изучен без построения полной теории вещественных чисел? Итак, что же такое континуум? Какие существуют вещественные числа? Когда и почему мы можем быть уверены, что действительно охватили совокупность *всех* вещественных чисел?

В элементах алгебры мы овладеваем множеством всех *рациональных* чисел, т. е. всех целых чисел и дробей, как положительных, так и отрицательных. Но очень скоро мы начинаем замечать, что этих чисел нам *не хватает*. Что это значит? Почему бы не ограничиться одними рациональными числами? Мы не можем согласиться на это потому, что среди рациональных чисел нет, например, числа $\sqrt{2}$, т. е. нет числа, квадрат которого равнялся бы числу 2. А почему такое число необходимо иметь? Потому хотя бы, что диагональ квадрата, сторона которого равна 1, имеет как раз длину $\sqrt{2}$; следовательно, если бы мы пожертвовали существованием такого числа, то нам пришлось бы мириться с тем, что длины отрезков, которые с такой простотой возникают в геометрии, могут не выражаться никакими числами. Ясно, что на такой базе измерительная геометрия развиваться не может. Значит, $\sqrt{2}$ должен найти себе место среди вещественных чисел; но среди рациональных чисел его нет; поэтому мы называем это число иррациональным. Однако этот $\sqrt{2}$ отнюдь не удовлетворяется тем, что мы признали его существующим; он немедленно начинает требовать, во-первых, чтобы среди рациональных чисел ему отвели совершенно определен-

ное место, т. е. точно указали, какие рациональные числа меньше его и какие больше (так, $\sqrt{2} > 1$ — диагональ квадрата больше его стороны); во-вторых, он требует, чтобы мы научились производить над ним действия, как над рациональными числами и совместно с ними (так, сумма стороны и диагонали квадрата равна $1 + \sqrt{2}$, что побуждает нас приписать смысл и этому числу, т. е. присоединить и его к совокупности вещественных чисел). Все эти требования нового числа вполне основательны и законны, и если мы в данный момент ничего на них не ответим, то это только потому, что мы сейчас же введем в наш мир еще много других новых чисел; все они без исключения будут предъявлять к нам те же самые претензии, и нам проще будет потом удовлетворить все эти претензии одновременно, чем заниматься в деталях каждым числом в отдельности.

За числом $\sqrt{2}$ в наш мир естественно и неизбежно входят все (положительные и отрицательные) квадратные корни из рациональных чисел, затем кубические корни и вообще все числа вида

$$r^{1/n}, \tag{1}$$

где r — любое положительное рациональное число, а n — любое целое число $\geqslant 2$. Но этим, как известно, дело не ограничивается; столь же конкретные запросы, как вышеприведенное соображение о диагонали квадрата, в целом ряде случаев заставляют нас вводить в качестве новых чисел корни тех или других алгебраических уравнений; это случается всякий раз, когда данное уравнение не имеет корней среди уже введенных нами чисел, между тем как мы не можем отказать этим корням в существовании, не лишая этим числовой характеристики какую-нибудь вполне конкретно определяемую величину.

Пойдем в этом направлении до конца. Назовем *алгебраическим числом* всякий корень (вещественный) уравнения вида $P(x) = 0$, где $P(x)$ — любой многочлен с целыми коэффициентами, и введем в наш мир все вещественные алгебраические числа; в частности, мы тем самым вводим и все числа вида (1), ибо число

$r^{1/n}$ определяется как корень уравнения $qx^n - p = 0$, где $\frac{p}{q} = r$ есть представление рационального числа r в виде простой дроби; в качестве еще более частного случая всякое рациональное число $r = \frac{p}{q}$ также принадлежит множеству алгебраических чисел, будучи корнем уравнения $qx - p = 0$.

Очень легко, как говорят, «упорядочить» весь этот мир алгебраических чисел, т. е. указать правило, позволяющее для любых двух алгебраических чисел определять, какое из них больше и какое меньше; несколько труднее, хотя тоже не очень сложно, определить правила всех обычных алгебраических действий над этими числами и показать, что результатами этих действий всегда являются снова алгебраические числа, так что — это очень важный момент! — алгебраические операции над алгебраическими числами всегда приводят к алгебраическим же числам и, значит, никогда не могут потребовать введения каких-либо новых чисел.

Может быть, мы могли бы теперь остановиться, считать построение мира вещественных чисел законченным и принять, следовательно, за континуум совокупность всех алгебраических чисел?

Хорошо известно, что мы этого сделать не можем, и известно, почему именно не можем. Если для многих алгебраических теорий построенными нами до сих пор числами можно удовлетвориться, то как раз для анализа такое ограничение является совершенно невозможным. Дело в том, что математический анализ к элементарным операциям алгебры уже на первых своих шагах присоединяет основную и важнейшую аналитическую операцию — *предельный переход*. Бывает очень много случаев, когда вполне конкретные соображения заставляют нас признать существование предела у той или другой последовательности чисел; мало того, предел этот встает перед нами как число, которое имеет конкретное и реальное значение и над которым нам в дальнейшем желательно производить алгебраические и аналитические операции. Если бы дело обстояло так, что для всякой последовательности алгебраических чи-

сел, которой мы считаем нужным приписать определенный предел, такой предел существовал действительно в той же области алгебраических чисел, то можно было бы смело допустить, что эта область и является континуумом, что ни в каких иных вещественных числах, кроме алгебраических, математический анализ не нуждается.

Но дело обстоит не так. Если мы возьмем круг радиуса 1 и будем вписывать в него правильные многоугольники, неограниченно увеличивая число их сторон, то периметры всех таких многоугольников выразятся алгебраическими числами. Предел этой числовой последовательности мы называем длиной окружности. Признать этот предел несуществующим означало бы запретить в геометрии говорить о длине окружности; вы легко можете представить себе, к какому крушению привел бы такой запрет не только геометрию, но и все другие точные науки, пользующиеся образом круга.

А между тем можно доказать, что среди алгебраических чисел такого предела нет. Где же выход из этого положения? Выход ясен. Мы вынуждены признать, что для целей математического анализа одних только алгебраических чисел недостаточно, что необходимо присоединить к ним еще новые вещественные числа; все такие неалгебраические вещественные числа называют обычно «трансцендентными»; построенное нами выше число (длину окружности радиуса 1) обозначают через 2π; значит, π есть число трансцендентное. Другим важным примером трансцендентного числа в анализе является известное число $e = 2{,}718...$, которое, как вы знаете, также порождается простым предельным переходом, исходя от рациональных чисел. Трансцендентность чисел e и π была доказана довольно поздно, во второй половине прошлого столетия. Однако необходимость введения трансцендентных чисел была показана несколько ранее, в середине прошлого столетия, на других, более простых, хотя и менее важных, примерах (французским математиком Лиувиллем).

Итак, наш континуум все еще не достроен. Как же нам следует поступить далее? Можем ли мы остановиться здесь и сказать, например, так: «Континуум есть

множество всех алгебраических чисел, к которым по мере надобности мы присоединяем (в качестве трансцендентных) числа, получающиеся (подобно e и π) из алгебраических путем предельного перехода, но сами не являющиеся алгебраическими»? Мы ставим этот вопрос потому, что ведь именно на такой базе (правда, обычно не высказываемой явно) строятся в большинстве случаев «упрощенные» курсы математического анализа, избегающие изложения полной теории иррациональных чисел.

Нет, конечно, так сказать нельзя, нельзя и остановиться там, где мы сейчас находимся, и на это имеется целый ряд весьма простых и убедительных причин. Прежде всего, континуум, будучи совокупностью *всех* вещественных чисел, должен определяться как некое твердое множество (по образцу, например, вышеустановленного определения совокупности всех алгебраических чисел), без оставления возможности впоследствии добавлять к нему все новые и новые числа. Далее, слова «по мере надобности» в нашем провизорном определении явно лишены точного содержания; если мы имеем последовательность алгебраических чисел, не имеющую алгебраического предела, то в вопросе о том, приписать ли ей трансцендентный предел или признать ее вовсе не имеющей предела, мы с формальной точки зрения в данный момент вправе поступать всякий раз по нашему произволу, и выбрать то или другое решение нас заставляют не формальные, а конкретно-реальные соображения, которые, сколь бы вески они ни были, не могут найти себе места в математическом определении. Формально мы можем отказать числу π в существовании; в данном случае это крайне невыгодно, но в других случаях подобный отказ может и не приводить ни к каким неудобствам; но ясно, что критерий, который хотел бы заставить нас вводить трансцендентное число «всякий раз, когда без этого неудобно обходиться», не есть и при любой перефразировке не может стать математическим критерием. Наконец, ниоткуда не следует, что введенными таким путем числами мы смогли бы удовлетвориться; ведь новые числа нам пришлось бы складывать, перемножать, подвергать предельным пере-

ходам (с числами, не допускающими этих операций, математическому анализу нечего делать); откуда мы можем иметь уверенность, что результаты всех операций будут вещественными числами нашего континуума? А если нет, то их снова придется добавить, и, следовательно, наш континуум не мог включать *всех* вещественных чисел.

Мы видим, таким образом, что на позиции, занимаемой нами в данный момент, удержаться нельзя; нельзя, построив в виде примера одно или два трансцендентных числа, сказать «и так далее» и на этом успокоиться, ибо, как мы только что убедились, никакого континуума мы таким образом не определим.

Итак, мы видим, что для полноценного обоснования математического анализа мы никак не можем обойтись без построения общей теории вещественных чисел, т. е. такой теории, которая не ограничивалась бы примерами конструкции новых чисел, а содержала в себе общие принципы такой конструкции, исчерпывающим образом характеризующие собой всю совокупность вещественных чисел.

Построение иррациональных чисел. В науке существует несколько различных теорий континуума. Однако все эти теории — и об этом важно всегда помнить — совершенно одинаково в идейном отношении подходят к своей задаче. По сравнению с этим их принципиальным единством то, что их между собою различает, может быть оцениваемо так же, как детали конструкции оцениваются в сравнении с архитектурным планом здания.

Все эти теории имеют своей целью, принимая в качестве первоначальной данности множество рациональных чисел, получить из него сразу всю совокупность всех вещественных чисел с помощью единого конструктивного принципа. Форма этого принципа различна в различных теориях; однако сходство между этими теориями еще далеко не заканчивается на указанном нами пункте. Дело в том, что выбор конструктивного принципа, ведущего к построению новых, иррациональных чисел, во всех теориях, несмотря на существенные формальные отличия, основан на одной и той же идее. Эта

идея состоит в том, что главная, доминирующая роль при конструкции новых чисел придается основной аналитической операции предельного перехода, все же другие встречающиеся способы порождения сводятся к ней, рассматриваются как ее частные случаи. Вы знаете, что действительно, например, квадратные корни из целых чисел являются пределами надлежаще выбранных последовательностей рациональных чисел («приближенных корней»); так же обстоит дело и в других случаях.

Ввиду сказанного, для более конкретного ознакомления с построением теории континуума нет необходимости рассматривать в деталях все эти различные способы обоснования, а вполне достаточно взять в качестве образца какой-нибудь один из них; все принципиально важное, что мы при этом увидим, будет в равной мере относиться и к другим теориям. Мы выбираем в дальнейшем теорию Дедекинда не потому, что она имеет какие-либо существенные преимущества перед другими, а лишь по той чисто внешней причине, что именно она принята в подавляющем большинстве наиболее распространенных руководств; таким образом, для читателя не представит затруднения отыскать такое пособие, где он сможет ознакомиться с деталями, пропущенными в нашем изложении.

1. Прежде чем приступить к введению иррациональных чисел, мы должны несколько более внимательно вглядеться в множество рациональных чисел, которое мы обозначим через R. Сначала обратим внимание на одно очень элементарное свойство этого множества: *между любыми двумя рациональными числами* r_1 *и* r_2 *всегда найдется третье*; проще всего мы в этом убеждаемся, замечая, что полусумма $\frac{r_1 + r_2}{2}$ есть всегда рациональное число, лежащее между r_1 и r_2; как следствие этого факта мы повторным его применением сейчас же устанавливаем, что между r_1 и r_2 всегда заключено бесконечное множество рациональных чисел.

2. Теперь постараемся внимательно всмотреться в ту картину, которая порождается при попытках найти или определить $\sqrt{2}$ (мы будем иметь в виду положительное значение радикала). Прежде всего мы ищем среди ра-

циональных чисел (никаких других для нас пока не существует!) такое число, квадрат которого равен числу 2, и легко обнаруживаем, что такого (рационального) числа не существует (мы не будем приводить здесь известного из курса средней школы арифметического доказательства этого факта). Это значит, что, какое бы рациональное число r мы ни выбрали, мы будем иметь либо $r^2 < 2$, либо $r^2 > 2$. Рассмотрим сперва одни только положительные рациональные числа. Они естественно разбиваются по только что рассмотренному признаку на два класса: класс A тех положительных рациональных чисел r_1, для которых $r_1^2 < 2$, и класс B тех положительных рациональных чисел r_2, для которых $r_2^2 > 2$. Так как r_1 и r_2 положительны, то из $r_1^2 < 2 < r_2^2$ следует $r_1 < r_2$, т. е. *каждое число класса A меньше каждого числа класса B*. Очевидно теперь, что в этом последнем заключении ничего не изменится, если мы присоединим к классу A нуль и все отрицательные (рациональные) числа; но тогда мы будем иметь уже разбиение всего множества R на два класса A и B, причем каждое число класса A меньше каждого числа класса B. Условимся называть *сечением* (точнее — сечением множества R) всякое разбиение множества R на два непустых [1]) класса, для которого имеет место это условие; мы пришли, таким образом, к некоторому определенному сечению множества R.

Сечения множества R мы можем строить самыми различными способами, в том числе и совершенно элементарными. Так, относя к классу A все рациональные числа $r_1 \leqslant 5$, а к классу B — все рациональные числа $r_2 > 5$, мы, очевидно, также получим некоторое определенное сечение множества R. Если обычным образом изображать числа точками на прямой линии, то всякое сечение, очевидно, изобразится некоторым разбиением (рациональных) точек прямой на два множества, из которых первое (A) целиком расположено влево от второго (B).

На первый взгляд может показаться, что все сечения множества R являют нам совершенно одинаковую

---

[1]) То есть содержащих хотя бы по одному числу.

картину, что два различных сечения отличаются друг от друга только тем местом, где они произведены, и что поэтому одно из них может быть переведено в другое простым сдвигом. Чрезвычайно важно, что *такое представление ошибочно,* что в самой структуре сечений возможно глубокое (и для нашей цели фундаментальное) различие.

В самом деле, в нашем последнем примере существует число (т. е. рациональное число, никаких других у нас покуда нет!), обладающее тем замечательным свойством, что все числа, меньшие его, принадлежат классу A, а все числа, бо́льшие его,— классу B; в нашем примере этим числом является, очевидно, число 5. Число, обладающее этим свойством, мы будем называть *рубежом* данного сечения; таким образом, в нашем последнем примере произведенное сечение имеет рубеж.

Напротив, *в нашем первом примере* (связанном с $\sqrt{2}$) *такого рубежа нет.* Докажем это.

Пусть рубежом служит (рациональное) число r; как для всякого (рационального) числа, мы должны иметь либо $r^2 < 2$, либо $r^2 > 2$; пусть для определенности $r^2 < 2$; по свойству рубежа для каждого $r' > r$ мы должны иметь $r'^2 > 2$. Но если $r < 1$, то, полагая $r' = 1$, мы уже приходим к противоречию; если же $r \geqslant 1$, то $r^2 \geqslant r$; поэтому, полагая $2 - r^2 = c\,(>0)$ и $r' = r + \frac{c}{4}$, мы имеем

$$r'^2 = r^2 + \frac{rc}{2} + \frac{c^2}{16} \leqslant r^2 + \frac{r^2c}{2} + \frac{c^2}{16} = 2 - \frac{7}{16}c^2 < 2,$$

что опять-таки приводит к противоречию, ибо $r' > r$.

Итак, все сечения множества R распадаются на два типа: имеющие рубеж и не имеющие рубежа; при этом надо иметь в виду, что

а) двух рубежей сечение иметь не может, ибо если r и r' — такие рубежи и если $r < r'$, то в силу пункта 1 существует такое число r'', что $r < r'' < r'$; но тогда по свойству рубежей из $r'' > r$ следует $r'' \in B$ [1]), а из $r'' < r'$ следует $r'' \in A$, что создает противоречие;

---

[1]) Запись $a \in M$ означает: a есть элемент множества M.

b) рубеж, если он существует, очевидно, есть либо наибольшее число класса A, либо наименьшее число класса B; если рубежа нет, то нет ни наибольшего числа в классе A, ни наименьшего числа в классе B;

c) каждое (рациональное) число r_0 служит рубежом двух различных сечений. В одном из них класс A состоит из чисел $r \leqslant r_0$ (а класс B — из чисел $r' > r_0$), в другом — класс A состоит из чисел $r < r_0$ (а класс B — из чисел $r' \geqslant r_0$);

d) распадение всей совокупности сечений множества R на два типа — имеющих рубеж и не имеющих рубежа,— очевидно, является внутренней структурной особенностью множества R; это есть факт, который остался бы в полной силе и в том случае, если бы мы и не помышляли никогда о введении каких-либо других чисел, кроме рациональных.

3. Теперь уже пример с числом $\sqrt{2}$ прямо подсказывает нам дальнейший образ действия. Картина ясна: с интуитивной, наглядной стороны перед нами — числовая прямая, вещественная ось, которую мы в определенном месте разрезали, причем оказалось, что в месте разреза находилась точка, которой никакое рациональное число не соответствует; мы не можем отказаться от мысли о существовании этой точки; без нее наша прямая и тот континуум, то множество чисел, которое пробегает переменная величина при своем изменении, утратили бы свою непрерывность, сплошность, т. е. именно ту черту, от которой континуум получил свое имя; говоря несколько грубо, нам прямо видно, что если бы мы заставили нашу переменную изменяться, принимая одни только рациональные значения, то ей пришлось бы при этом от одного значения к другому перепрыгивать через зияющие пустоты; со стороны практической, как мы уже раньше об этом говорили, все прикладные науки (и прежде всего геометрия) вынуждены были бы испытывать весьма значительные неудобства, если бы мы примирились с отсутствием какого бы то ни было рубежа между нашими двумя классами. Поэтому мы, побуждаемые как требованиями нашего наглядного представления, так и вескими практи-

ческими соображениями, вводим в наш числовой мир новое число $\sqrt{2}$, которое по определению служит рубежом нашего сечения и которое мы называем *иррациональным числом.*

Но выбранное нами сечение ничем принципиально не отличается от любого другого сечения множества R того же типа, т. е. не имеющего (рационального) рубежа. В порядке построения общей теории мы поэтому естественно распространяем наше определение на случай любого такого сечения. Каждому сечению множества R, не имеющему рационального рубежа, мы ставим в соответствие некоторое новое, иррациональное число, которое по определению является его рубежом.

Таким образом, с помощью этого единого принципа мы сразу определяем все множество иррациональных чисел. Вместе с существовавшими уже ранее рациональными числами они образуют множество всех вещественных чисел, или континуум, который является теперь полностью определенным.

Теория континуума. Однако введенным нами принципом построения иррациональных чисел теория континуума, разумеется, не исчерпывается; наоборот, здесь она, в сущности, только начинается; программа работ, которые мы должны выполнить, прежде чем мы сможем говорить о действительно завершенной теории континуума, еще чрезвычайно обширна. Мы должны, прежде всего, *упорядочить* наш континуум, т. е. точно определить, при каких условиях данное вещественное число мы будем считать бо́льшим или меньшим другого данного вещественного числа. Далее, мы должны определить действия над вещественными числами; ведь до сих пор мы не имеем, например, ни малейшего представления о том, что значит сложить между собою числа 1 и $\sqrt{2}$; далее, мы должны тщательным образом убедиться, что эти действия обладают всеми теми свойствами, к которым мы привыкли в области рациональных чисел; так, независимость суммы от порядка слагаемых (коммутативный, или переместительный, закон сложения) есть теорема, которую мы обязаны заново обосновать для любых вещественных чисел; наконец, мы

должны найти способ убедиться, что определенный нами континуум действительно отвечает всем тем запросам практики и нашего наглядного представления, для удовлетворения которых он был нами построен. Разумеется, детальное выполнение этой программы в рамках наших лекций совершенно невозможно; да оно было бы и в высшей степени скучным; мы в дальнейшем коснемся лишь некоторых, принципиально наиболее важных моментов этой программы.

Прежде всего, очень легко упорядочить наш континуум. Пусть мы имеем два вещественных числа α_1 и α_2; требуется определить, какое из них больше и какое меньше другого. Если оба числа рациональны, то этот вопрос решается в арифметике и здесь нас занимать не должен; если, например, α_1 иррационально, а α_2 рационально, то вопрос также немедленно решается: α_1 есть рубеж некоторого сечения множества R; в соответствии с самим определением рубежа мы полагаем $\alpha_2 < \alpha_1$ или $\alpha_2 > \alpha_1$ в зависимости от того, принадлежит ли рациональное число α_2 классу A или классу B в этом сечении. Пусть, наконец, α_1 и α_2 оба иррациональны; так как эти числа между собою различны, то различны между собою и те два сечения, рубежами которых они служат; в частности, различны между собою левые классы A_1 и A_2 этих двух сечений; это значит, что в каком-нибудь из этих множеств, например в A_2, найдется такое рациональное число r, которого нет в A_1; из $r \in A_2$ следует $r < \alpha_2$; из $r \overline{\in} A_1$ [1]) следует $r \in B_1$, и, следовательно, $r > \alpha_1$; таким образом, существует такое рациональное число r, что $\alpha_1 < r < \alpha_2$; в этом случае мы условимся считать $\alpha_1 < \alpha_2$; если бы, наоборот, нашлось такое r', что $\alpha_2 < r' < \alpha_1$, то мы считали бы $\alpha_2 < \alpha_1$; одно из двух, как мы сейчас убедились, обязательно должно иметь место, и этим, очевидно, упорядочение континуума завершено.

Впрочем, завершено только определение, свойства его еще должны быть доказаны; надо убедиться, что $\alpha_1 < \alpha_2$ несовместимо с $\alpha_1 > \alpha_2$; что из $\alpha_1 < \alpha_2$ вытекает $\alpha_2 > \alpha_1$, и обратно; наконец, что из $\alpha_1 < \alpha_2$ и

---

[1]) Запись $a \overline{\in} M$ означает: a не есть элемент множества M.

$\alpha_2 < \alpha_3$ вытекает $\alpha_1 < \alpha_3$,— словом, что неравенства между вещественными числами подчиняются тем же основным законам, что и неравенства между рациональными числами. Вы легко справитесь сами с доказательствами всех этих предложений.

Последнее проведенное нами рассуждение показывает, между прочим, что между любыми двумя иррациональными числами найдется всегда рациональное число; мы знаем, что то же самое остается верным, если оба данных числа рациональны. Пусть теперь число r рационально, число α иррационально, и пусть для определенности $r < \alpha$; покажем, что и в этом случае между r и α найдутся рациональные числа. Число α есть рубеж сечения (A, B) множества R, причем из $r < \alpha$ вытекает $r \in A$; но в сечении с иррациональным рубежом класс A не содержит наибольшего числа, поэтому существует рациональное число $r' > r$, которое, так же как и r, принадлежит классу A и потому меньше, чем α; таким образом, $r < r' < \alpha$, что и требовалось доказать.

Итак, мы убедились, что между двумя любыми вещественными числами найдется рациональное число, а значит, очевидно, и бесчисленное множество рациональных чисел; это важное свойство множества R выражают обычно, говоря, что R *всюду плотно* (в континууме). Легко убедиться, что и множество иррациональных чисел всюду плотно; проще всего сделать это, рассматривая совокупность всех чисел вида $r\sqrt{2}$, где r — любое рациональное число, отличное от нуля; все такие числа иррациональны, и уже они одни образуют всюду плотное множество; впрочем, говоря строго, выражение $r\sqrt{2}$ получает точный смысл только после того, как определены действия над вещественными числами, к чему мы теперь и переходим.

У нас нет оснований рассматривать этот вопрос во всех деталях, так как для выяснения всей идейной стороны этого определения нам достаточно будет внимательно разобрать один пример. Пусть α_1 и α_2 — два вещественных числа, и пусть мы хотим определить их сумму $\alpha_1 + \alpha_2$. Будут ли числа α_1 и α_2 рациональными или иррациональными, во всяком случае они оба служат рубежами некоторых сечений множества R, кото-

рые мы обозначим соответственно через (A_1, B_1) и (A_2, B_2). Пусть, далее, a_1, b_1, a_2, b_2 означают любые числа, принадлежащие соответственно множествам A_1, B_1, A_2, B_2. Очевидно, что всякое рациональное число вида $a_1 + a_2$ меньше любого рационального числа вида $b_1 + b_2$. Мы теперь установим существование и единственность такого вещественного числа α, что при любых a_1, a_2, b_1, b_2 (принадлежащих, конечно, соответствующим множествам)

$$a_1 + a_2 \leqslant \alpha \leqslant b_1 + b_2.$$

Это число α мы естественным образом и определим как сумму чисел α_1 и α_2, существование и единственность которой будут тем самым установлены для всех случаев.

Рассмотрим сечение (A, B) множества R, определяемое следующим образом: если рациональное число r меньше всех чисел вида $b_1 + b_2$, мы отнесем его к классу A; в противном случае — к классу B (вы убедитесь очень легко, что так определенное разбиение множества R действительно есть сечение). Рубеж этого сечения обозначим через α. Очевидно, что все числа вида $a_1 + a_2$ принадлежат классу A, а все числа вида $b_1 + b_2$ — классу B; поэтому

$$a_1 + a_2 \leqslant \alpha \leqslant b_1 + b_2, \tag{2}$$

т. е. число α действительно удовлетворяет поставленным условиям. Нам остается установить его единственность.

С этой целью убедимся сначала, что числа a_1 и b_1 мы можем выбрать с таким расчетом, чтобы их разность $b_1 - a_1$ была как угодно малой. В самом деле, пусть a — любое рациональное число класса A_1 и пусть c — как угодно малое положительное рациональное число; в последовательности рациональных чисел

$$a, a + c, a + 2c, \ldots, a + nc, \ldots$$

первый член принадлежит классу A_1; в дальнейшем, вообще говоря, еще несколько членов будут принадлежать классу A_1; но с некоторого места все последующие члены уже будут принадлежать классу B_1, так как $a + nc$

безгранично возрастает вместе с n. Поэтому найдется такое целое число k, что

$$a + kc = a_1 \in A_1, \quad a + (k+1)c = b_1 \in B_1,$$

откуда разность $b_1 - a_1 = c$ и, следовательно, сколь угодно мала; этим наше утверждение и доказано. По такой же причине, разумеется, и числа a_2, b_2, а значит, и числа $a_1 + a_2$, $b_1 + b_2$ могут быть выбраны сколь угодно близкими друг к другу.

Пусть теперь, если это возможно, существуют два вещественных числа α и α', удовлетворяющих всем неравенствам вида (2), и пусть для определенности $\alpha < \alpha'$. Как мы знаем, существует такая пара рациональных чисел r, r', что

$$\alpha < r < r' < \alpha';$$

вместе с неравенствами (2) эти соотношения показывают, что любое число вида $a_1 + a_2$ отстоит от любого числа вида $b_1 + b_2$ на величину, превышающую постоянное число $r' - r$. Это же, очевидно, противоречит только что полученному нами результату. Таким образом, доказана и единственность числа α.

Приведенное нами определение сложения по своей форме удобно в том отношении, что позволяет непосредственно распространить все основные законы, которым подчиняется сложение рациональных чисел, на сложение любых вещественных чисел. Попытайтесь проделать это хотя бы для переместительного закона, и вы немедленно увидите, как просто все это выходит.

Как мы уже сказали, мы здесь не будем останавливаться на определении других действий и доказательстве управляющих ими законов. Заметим только, что умножение лучше всего определять аналогично сложению, а вычитание и деление — как обратные действия.

Мы же теперь обратимся к последнему крайне важному в этом круге идей вопросу: как убедиться, что определенный нами континуум действительно являет собой ту непрерывность, сплошность, которая необходима ему как базе математического анализа и отсутствие которой у множества R заставило нас в свое время прибегнуть к введению иррациональных чисел?

Чтобы ответить на этот вопрос, вспомним, что́ дало нам повод говорить об отсутствии такой непрерывности у множества R. Это был тот факт, что среди сечений множества R оказались не имеющие рубежа, принадлежащего этому множеству R. Поэтому, если мы покажем, что для континуума такой факт никогда не может иметь места, т. е. что *всякое сечение континуума имеет рубеж, принадлежащий этому самому континууму,* то мы сможем считать себя удовлетворенными и надеяться, что континуум, построенный нами в качестве базы математического анализа, будет отвечать предъявляемым к нему запросам.

Заметим во избежание недоразумений, что только что формулированное нами предложение не вытекает непосредственно из предыдущего, ибо до сих пор мы всегда говорили о сечениях множества R, теперь же впервые говорим о сечении континуума; формально определение сечения мы при этом мыслим себе неизменным.

Докажем теперь наше утверждение. Пусть (A, B) есть любое сечение континуума; всякое рациональное число, как и вообще всякое вещественное число, принадлежит либо классу A, либо классу B; таким образом, сечение (A, B) континуума непосредственно индуцирует некоторое совершенно определенное сечение (A', B') множества R. Пусть α — вещественное число, служащее рубежом сечения (A', B'); убедимся, что оно же служит рубежом сечения (A, B), чем наше утверждение и будет доказано.

Нам предстоит показать, что всякое вещественное число $\alpha_1 < \alpha$ принадлежит классу A, а всякое вещественное число $\alpha_2 > \alpha$ — классу B; из соображений симметрии достаточно доказать первое утверждение. Пусть r — рациональное число, лежащее между α_1 и α; так как $r < \alpha$, то $r \in A' \subset A$ [1]); а так как $\alpha_1 < r$, то отсюда вытекает, что и $\alpha_1 \in A$, что и требовалось доказать.

Основные леммы. Логический фундамент математического анализа, таким образом, построен. Последова-

[1]) Запись $A \subset B$ означает: множество A содержится в множестве B.

тельно воздвигая на этой базе основы анализа, мы, разумеется, бываем вынуждены очень часто ссылаться на заложенный фундамент, т. е. обращаться непосредственно к установленному нами определению вещественного числа. Это сопряжено с известными неудобствами, так как построение и исследование необходимых при этом сечений обычно бывает в достаточной мере громоздким.

Тот путь, которым наука находит выход из этого затруднения, в высшей степени поучителен, так как его можно считать типичным для всех подобного рода логических ситуаций, часто встречающихся в математической науке. В ходе развития математического анализа было замечено, что хотя непосредственно применять в рассуждениях определение вещественных чисел с помощью сечений приходится очень часто, многие из этих применений в формальном отношении весьма похожи друг на друга, так что фактически почти вся совокупность этих применений совершается по трем-четырем формальным схемам (наполняемым, конечно, всякий раз особым содержанием). Но если создалось такое положение, то было бы, конечно, очень неэкономно и весьма затруднило бы как построение, так и усвоение данной дисциплины, если бы мы стали десятки раз повторять одну и ту же логическую конструкцию, наполняя ее только всякий раз новым предметным содержанием. Математическая наука уже давно — и, конечно, с полным основанием — усвоила обычай во всех подобного рода случаях явно формулировать такую логическую схему в виде некоторого вспомогательного предложения (леммы), с тем чтобы, раз навсегда доказав эту лемму, в дальнейшем уже иметь возможность не повторять всякий раз той формальной конструкции, которая лежит в ее основе, а просто ссылаться на готовую лемму. Доказав три-четыре таких вспомогательных предложения, мы получаем возможность во всем дальнейшем уже почти никогда не возвращаться к конструкции сечений, всякий раз заменяя такую конструкцию ссылкой на одну из основных лемм, составляющих как бы небольшую группу мостов, соединяющих математический анализ с его логической базой. Само собой

разумеется, что выбор этих основных лемм может быть различным в различных изложениях; однако во всех случаях можно рекомендовать не жалеть времени и сил на возможно тщательное усвоение возможно большего числа таких лемм, ибо каждая из них имеет своим назначением существенное облегчение работы в будущем, и поэтому затраченные на овладение ею усилия, без всяких сомнений, не пропадут даром.

Примерный тип формулировки и доказательства такого рода лемм мы покажем теперь на нескольких примерах.

1. Лемма о монотонной последовательности. Последовательность вещественных чисел

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots \qquad (3)$$

называется *монотонной*, если либо $\alpha_n \leqslant \alpha_{n+1}$ для любого n, либо $\alpha_n \geqslant \alpha_{n+1}$ для любого n; в первом случае мы говорим, точнее, о *монотонно неубывающей,* во втором — о *монотонно невозрастающей* последовательности. Последовательность (3) называется *ограниченной,* если существует такое положительное число c, что

$$|\alpha_n| < c \quad (n = 1, 2, \ldots).$$

Лемма 1. *Всякая монотонная ограниченная последовательность имеет предел* [1]).

Пусть для определенности $\alpha_n \leqslant \alpha_{n+1}$ и $|\alpha_n| < c$ $(n = 1, 2, \ldots)$. Разобьем континуум на два множества A и B, помещая в B всякое вещественное число, превосходящее *все* числа α_n (так что, например, число c входит в множество B), а в A — все остальные числа (в частности, например, все α_n); очевидно, что это разбиение континуума есть сечение. Пусть α — рубеж этого сечения. Покажем, что $\lim\limits_{n\to\infty} \alpha_n = \alpha$, чем лемма 1 и будет доказана.

Прежде всего заметим, что для любого n

$$\alpha_n \leqslant \alpha;$$

в самом деле, при $\alpha_n > \alpha$ мы по определению рубежа

---

[1]) Определение предела последовательности см. на стр. 37.— *Прим. ред.*

имели бы $\alpha_n \in B$, что противоречит определению множества B. Если, вопреки нашему утверждению, α не есть предел последовательности (3), то существует такое постоянное положительное число ε, что для бесконечного множества чисел n

$$\alpha - \alpha_n > \varepsilon,$$

откуда $\alpha_n < \alpha - \varepsilon$; но в силу монотонности последовательности (3) это неравенство, выполняясь для бесконечного множества значений n, должно выполняться для *всех* n; по определению множества B отсюда следует $\alpha - \varepsilon \in B$, тогда как ввиду $\alpha - \varepsilon < \alpha$ число $\alpha - \varepsilon$ должно принадлежать классу A [ибо α есть рубеж сечения (A, B)]. Полученное противоречие доказывает наше утверждение.

Разумеется, в предпосылках леммы монотонность не может быть отброшена; из одной только ограниченности последовательности (3) не вытекает существование предела, как показывает пример $\alpha_n = (-1)^n$.

Лемма о монотонной последовательности находит себе весьма многочисленные применения не только в анализе, но даже и в элементарной математике; в последнем случае ее часто вводят в качестве «аксиомы», иногда не учитывая того, что при имеющемся в распоряжении учащихся весьма скудном запасе вещественных чисел эта «аксиома» просто неверна; впрочем, этот дефект приходится наблюдать не только в элементарной математике, но и в «упрощенных» курсах математического анализа.

В частности, существование длины окружности (т. е. предела периметров вписанных правильных многоугольников при безграничном увеличении числа сторон) в геометрии и существование числа $e = \lim\limits_{n \to \infty}\left(1 + \frac{1}{n}\right)^n$ в основах анализа доказываются проще всего ссылкой на эту лемму.

Наряду с леммой 1, характеризующей основную черту монотонных ограниченных последовательностей, не меньшее значение имеет для анализа и аналогичное предложение, касающееся монотонных функций непрерывно изменяющейся величины.

Лемма 1′. *Если величина x, возрастая, стремится к a и если функция $f(x)$ ограничена и монотонна в некотором отрезке, имеющем точку a своим правым концом, то $f(x)$ стремится к определенному пределу при $x \to a$.*

При этом ограниченность функции $f(x)$ внутри отрезка с концами $a-\varepsilon$ и a $(\varepsilon > 0)$ означает существование такого числа c, что

$$|f(x)| < c \qquad (a-\varepsilon < x < a);$$

монотонность же ее означает, что отношение $\frac{f(x_1)-f(x_2)}{x_1-x_2}$ либо $\geqslant 0$ для любой пары x_1, x_2, $x_1 \neq x_2$, чисел этого отрезка, либо $\leqslant 0$ для любой такой пары.

Доказательство леммы 1′ очень просто проводится на основании леммы 1. Пусть для определенности функция $f(x)$ неубывающая, т. е. $f(x_2) \geqslant f(x_1)$ при $a > x_2 > x_1 > a-\varepsilon$. Так как при $n > \frac{1}{\delta}$

$$a-\delta < a-\frac{1}{n} < a,$$

то возрастающая числовая последовательность $a-\frac{1}{n}$ при $n > \frac{1}{\varepsilon}$ расположена внутри отрезка с концами $a-\varepsilon$ и a и имеет пределом a; соответствующая последовательность чисел $f\left(a-\frac{1}{n}\right)$, очевидно, ограниченная и неубывающая; поэтому в силу леммы 1 существует предел $\lim\limits_{n\to\infty} f\left(a-\frac{1}{n}\right)=b$. Если число x достаточно близко к a, то найдется такое $n=n(x)$, что

$$a-\frac{1}{n} \leqslant x < a-\frac{1}{n+1},$$

а значит, и

$$f\left(a-\frac{1}{n}\right) \leqslant f(x) \leqslant f\left(a-\frac{1}{n+1}\right) \tag{4}$$

в силу монотонности функции f. Но при $x \to a$, очевидно, $n \to \infty$, вследствие чего левая и правая части

неравенств (4) имеют своим пределом b, откуда следует, что и $f(x) \to b$ при $x \to a$; этим лемма 1′ доказана.

2. Лемма о стягивающейся последовательности отрезков. Условимся закрытый отрезок с концами a и b, т. е. совокупность всех вещественных чисел x, удовлетворяющих неравенствам $a \leqslant x \leqslant b$, обозначать через $[a, b]$, а открытый отрезок с теми же концами, т. е. совокупность всех вещественных чисел x, удовлетворяющих неравенствам $a < x < b$, обозначать через (a, b). Последовательность отрезков

$$[a_1, b_1],\ [a_2, b_2],\ \ldots,\ [a_n, b_n],\ \ldots \qquad (5)$$

мы будем называть *стягивающейся*, если она подчинена следующим двум требованиям:

1°. $a_n \leqslant a_{n+1} < b_{n+1} \leqslant b_n \quad (n = 1, 2, \ldots)$, т. е. каждый последующий отрезок целиком содержится в предыдущем;

2°. $\lim\limits_{n\to\infty}(b_n - a_n) = 0$, т. е. длины отрезков при безграничном возрастании их номеров стремятся к нулю.

Лемма 2. *Если (5) — стягивающаяся последовательность отрезков, то существует единственное вещественное число α, принадлежащее всем этим отрезкам.*

Можно было бы, конечно, провести доказательство этой (часто очень полезной) леммы, конструируя соответственные сечения; однако значительно проще — сослаться на лемму 1. В самом деле, в силу условия 1° последовательность

$$a_1, a_2, \ldots, a_n, \ldots$$

монотонна и ограничена (последнее видно из того, что $a_n < b_1$ для всех n); в силу леммы 1 она имеет предел; положим

$$\lim_{n\to\infty} a_n = \alpha.$$

Так как при любом k

$$a_n \leqslant b_k \quad (n = 1, 2, \ldots),$$

то и

$$\alpha \leqslant b_k \quad (k = 1, 2, \ldots),$$

так что

$$a_n \leqslant \alpha \leqslant b_n \quad (n = 1, 2, \ldots). \qquad (6)$$

Таким образом, число α принадлежит каждому из отрезков $[a_n, b_n]$ и, следовательно, удовлетворяет условиям леммы. Единственность его вытекает из того, что при наличии двух чисел α и β, удовлетворяющих неравенствам (6), мы имели бы (предполагая для определенности $\alpha < \beta$)

$$a_n \leqslant \alpha < \beta \leqslant b_n \quad (n = 1, 2, \ldots),$$

откуда

$$b_n - a_n \geqslant \beta - \alpha \quad (n = 1, 2, \ldots),$$

что противоречит свойству 2° стягивающейся последовательности. Таким образом, лемма 2 доказана.

Стоит заметить, что для леммы 2 весьма существенно причисление к каждому отрезку его концов a_n, b_n; если бы мы отказались от этого и вместо $[a_n, b_n]$ рассматривали открытый отрезок $a_n < x < b_n$, то лемма была бы неверна; так, последовательность открытых отрезков $0 < x < \frac{1}{n}$ $(n = 1, 2, \ldots)$ не имеет общей всем отрезкам точки.

3. Лемма Гейне — Бореля. Это вспомогательное предложение более недавнего происхождения также часто бывает очень удобным орудием доказательства теорем анализа.

Условимся говорить, что (вообще говоря, бесконечная) система M отрезков *покрывает* отрезок (закрытый) $[a, b]$, если каждая точка этого последнего лежит внутри по меньшей мере одного из отрезков системы M.

Лемма 3. *Если система M покрывает отрезок $[a, b]$, то из нее можно выделить конечную систему отрезков M', также покрывающую отрезок $[a, b]$.*

В самом деле, если отрезок $\Delta_1 = [a, b]$ не допускает требуемого леммой конечного покрытия, то, разделив его пополам, мы можем утверждать, что по меньшей мере одна из двух половин также не допускает конечного покрытия отрезками системы M (так как, очевидно, если бы обе его допускали, то допускал бы его и весь отрезок Δ_1). Обозначим эту половину через Δ_2

(если обе половины не допускают такого конечного покрытия, то Δ_2 может означать любую из них) и разделим ее опять пополам; снова мы можем утверждать, что по меньшей мере одна из двух половин (обозначим ее через Δ_3) не допускает такого конечного покрытия. Этот процесс мы можем продолжать безгранично, причем отрезки Δ_1, Δ_2, Δ_3, ..., очевидно, образуют стягивающуюся систему отрезков. На основании леммы 2 существует единственная точка α, принадлежащая всем этим отрезкам. Пусть Δ — отрезок системы M, содержащий внутри себя точку α; так как при $n \to \infty$ длина отрезка Δ_n стремится к нулю, причем $\alpha \in \Delta_n$, то $\Delta_n \subset \Delta$ для достаточно больших n, и мы приходим к противоречию: отрезок Δ_n, по своему определению не допускающий конечного покрытия, покрывается одним отрезком Δ системы M. Этим противоречием и доказывается лемма 3.

Теперь мы не только научились строить фундамент математического анализа, но, доказав три важнейших вспомогательных предложения, подготовили этот фундамент к наиболее удобному применению его в процессе дальнейшей конструкции основного здания. Какими методами воздвигается это здание, каковы основные понятия, идеи и способы рассуждения, употребительные при его построении,— все это вы узнаете из следующих лекций.

Л Е К Ц И Я II

ПРЕДЕЛЫ

Что такое предел.— Различные типы стремления к пределу.— Предел постоянной величины.— Бесконечно малые и бесконечно большие.— Критерий Коши.— Замечание об основных теоремах.— Частичные пределы; верхний предел и нижний предел.— Пределы функций нескольких переменных.

Что такое предел. Понятие предела, наряду с понятием функциональной зависимости, есть одна из важнейших концепций математического анализа. Вы, конечно, знаете много теорем о пределах; и тем не менее нам придется со всей основательностью заняться этим понятием, чтобы отчасти уточнить, а отчасти дополнить, углубить и расширить ту его концепцию, к которой вы привыкли.

Займемся прежде всего вопросом о том, что означает фраза «переменная величина x (в данном явлении или процессе) стремится к числу a (или, что то же, имеет своим пределом число a)» — фраза, которую мы символически записываем одним из двух способов: $x \to a$ или $\lim x = a$. Если мы хотим при определении предела соблюсти все формальные требования (а без этого вообще нельзя строить математическую науку), то мы сейчас же сталкиваемся со своеобразным и характерным затруднением. Дело в том, что в точном определении понятия предела, конечно, не может быть места таким терминам, как «явление» или «процесс», математический смысл которых совершенно неясен. А между тем без этих (или равнозначных им) терминов формулировать обычное, знакомое вам определение понятия предела очень затруднительно. Мы ведь обычно гово-

рим так: какова бы ни была окрестность U числа a [1]), все значения величины x, *начиная с некоторого момента данного процесса*, принадлежат окрестности U [2]). И вот теперь мы должны найти способ так формулировать это определение, чтобы оно не содержало никаких терминов, точный математический смысл которых не был бы определен заранее. Надо прямо сказать, что задача эта вряд ли удовлетворительно разрешима. Современные математики предпочитают вовсе отказаться от ее решения и считают запись $\lim x = a$ лишенной математического содержания. Это не значит, конечно, что они вовсе отказываются от понятия предела. Как же они выходят из этого затруднения?

Дело в том, что фактически мы в анализе всегда встречаемся с таким положением вещей, когда к пределу стремится некоторая *функция* y при некотором определенном поведении независимой переменной x, т. е., например, «y имеет пределом b, когда x стремится к a» или «a_n стремится к α, когда n безгранично возрастает» и т. п. Такая фраза (и соответствующая ей запись $y \underset{x\to a}{\to} b$, или $\lim\limits_{x\to a} y = b$, или $a_n \underset{n\to\infty}{\to} \alpha$) имеет уже совершенно определенный смысл. Так, выражение $\lim\limits_{x\to a} y = b$ означает следующее: *какова бы ни была окрестность V числа b, существует такая окрестность U числа a, что $y \in V$ всякий раз, когда $x \in U$, за исключением, быть может, $x = a$* [3]). Вы видите, что это безукоризненно четкое определение понятия предела действительно не нуждается в представлении о процессе. Оно просто говорит, что y сколь угодно близко к b при x, достаточно

---

[1]) Окрестность числа a есть любой открытый отрезок, содержащий число a внутри себя.

[2]) Конечно, иногда редактируют эту фразу и иначе; так, говорят: сколь бы мало ни было положительное число ε, величина $x - a$ с *течением данного процесса* становится и остается по абсолютному значению меньше, чем ε. Как бы то ни было, во всех распространенных определениях фигурирует это упоминание о процессе, без которого самое понятие предела кажется лишенным определенности.

[3]) Функция y может быть не определена при $x = a$; может также случиться, что y при $x = a$ имеет определенное значение, но это значение отлично от b.

близком к a[1]). Конечно, для того, кто не свыкся еще с типическими чертами математического формализма, может показаться странным: как же, в самом деле, фраза (А) «предел y равен b при условии, что предел x равен a» может иметь определенный смысл, в то время как *само условие* «предел x равен a» не имеет смысла? На самом же деле в таком положении вещей нет ничего не только неприемлемого, но даже и необычного. Фраза (А) определяется как нечто монолитное, целое, и вовсе не обязательно, чтобы отдельные выхваченные нами ее части также имели смысл, если только всей этой фразе, взятой целиком, приписывается точно определенное содержание.

Само собой понятно, что данное нами определение понятия предела включает в себя оба наиболее важных для анализа случая — предел последовательности $\lim\limits_{n\to\infty} a_n$[2]) и предел функции от непрерывно изменяющейся величины $\lim\limits_{x\to a} y$, где a есть либо вещественное число, либо $+\infty$, либо $-\infty$; точно так же предел самой функции y может здесь пониматься в любом из этих трех смыслов.

Различные типы стремления к пределу. Пусть $y = f(x)$ есть функция от x, имеющая пределом число b, когда x стремится к a, причем a также число; предположим для простоты, что все значения величины x больше a; символически это выражают очень удобной записью: $x \to a + 0$ (вместо $x \to a$, $x > a$); таким образом, мы имеем

$$\lim_{x\to a+0} f(x) = b. \tag{1}$$

---

[1]) При этом случаи $x \to +\infty$, $x \to -\infty$, $y \to +\infty$, $y \to -\infty$ не требуют специальных определений, если под окрестностью $+\infty$ $(-\infty)$ понимать любое множество чисел, бо́льших (меньших) некоторого определенного числа.

[2]) Число b называется пределом последовательности $a_1, a_2, \ldots, a_n, \ldots$, если для всякого $\varepsilon > 0$ существует такой номер N, что для всех $n > N$ будет $|a_n - b| < \varepsilon$. Это определение будет частным случаем общего определения предела, данного выше, если под окрестностью символа ∞ понимать множество всех натуральных чисел, превосходящих некоторое число N. — *Прим. ред.*

Допустим сначала, что для всех $x > a$, достаточно близких к a, мы имеем $f(x) > b$; тогда функция $f(x)$ имеет вблизи точки a, в обычном геометрическом изображении, один из двух видов, иллюстрируемых на рис. 1 и 2. В случае рис. 1 $f(x)$ тем ближе к b, чем x

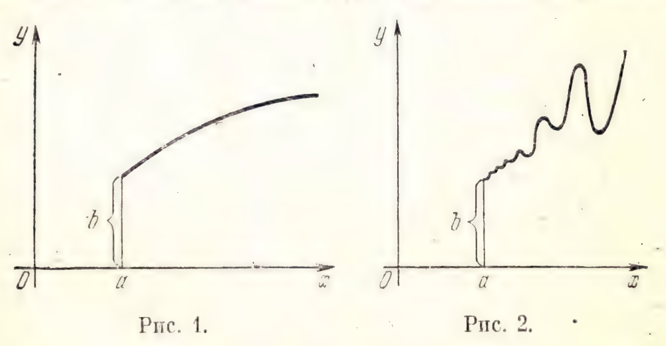


Рис. 1. Рис. 2.

ближе к a; $y \to b$, монотонно убывая при $x \to a + 0$. В случае рис. 2 все это не так: $y \to b$, то приближаясь, то удаляясь от своего предела; при $x \to a + 0$ функция $f(x)$ изменяется не монотонно: она то возрастает, то убывает. Разумеется, что все эти существенные различия не препятствуют нам аналитически выражать обе картины одним и тем же соотношением (1), описывающим основной факт, общий обоим случаям: y как угодно близко к b, если x достаточно близко к a.

Случай, когда для всех $x > a$, достаточно близких к a, мы имеем $y < b$, совершенно аналогичен предыдущему; он охватывается картинами, изображенными на рис. 3 и 4, и, очевидно, не нуждается в пояснениях.

Наконец, может быть случай, когда вправо от точки a и в любой близости к этой точке функция $y = f(x)$ принимает как значения, бо́льшие чем b, так и значения, меньшие чем b (рис. 5). В этом случае стремление величины y к пределу b при $x \to a + 0$ не может, очевидно, быть монотонным. При этом, если функция $f(x)$ непрерывна (как мы неявно предполагали на всех рисунках), то она необходимо становится равной своему предельному значению b в бесчисленном множестве точек, лежащих вправо от a и сколь угодно близких к a;

говоря геометрически, график функции $f(x)$ в соседстве точки a бесчисленное множество раз пересекает прямую $y = b$.

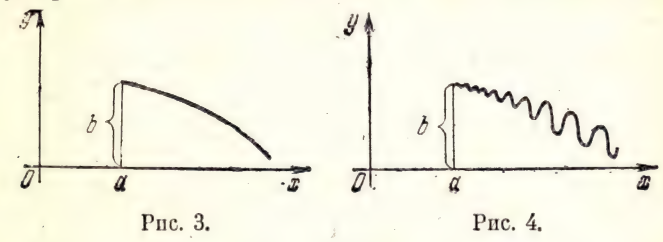


Рис. 3. Рис. 4.

Этими случаями исчерпываются, очевидно, все возможные типы непрерывно изменяющихся величин, стремящихся к b при $x \to a + 0$. Мы не будем отдельно рассматривать случай $x \to a - 0$, когда все значения величины $x < a$; соответствующие картины получаются, очевидно, зеркальным отражением рис. 1—5 в прямой $x = a$.

Если, как мы предположили, $y \to b$ при $x \to a$, то, конечно, мы имеем как $\underset{x \to a+0}{y \to b}$, так и $\underset{x \to a-0}{y \to b}$; при этом любой из пяти упомянутых нами типов поведения функции y вправо от точки a может сочетаться с любым из пяти аналогичных типов ее поведения влево от точки a, так что мы получаем всего 25 различных типов поведения функции y, стремящейся к b при $x \to a$.

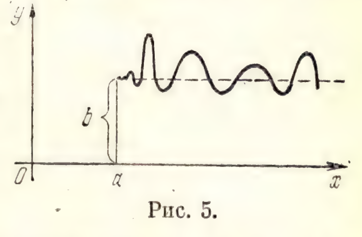


Рис. 5.

Заметим еще, что случаи, изображенные на рис. 1, 2 и 3, 4, удобно иногда записывать соответственно в виде

$$\underset{x \to a+0}{y \to b + 0} \text{ и } \underset{x \to a+0}{y \to b - 0}.$$

До сих пор мы предполагали a и b числами. Нетрудно перечислить все возможные типы поведения функции y и в тех случаях, когда одна из этих букв (или обе буквы) означает $+\infty$ или $-\infty$. Вы сами без труда их найдете; для примера укажем, что в случае, когда

$a = +\infty$, а b — число, возможны типы, изображенные на рис. 1′—5′.

Функция $f(x)$ монотонна в случае рис. 1′ и 3′ и не монотонна в остальных случаях; в случае рис. 5′ функция $f(x)$ (если она непрерывна) становится равной

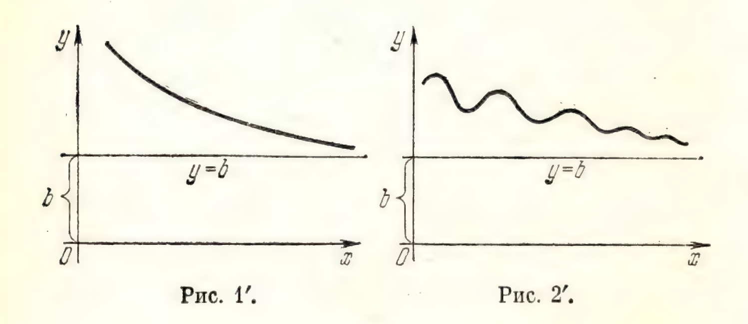


Рис. 1′. Рис. 2′.

своему предельному значению b бесчисленное множество раз, и притом для сколь угодно больших значений x.

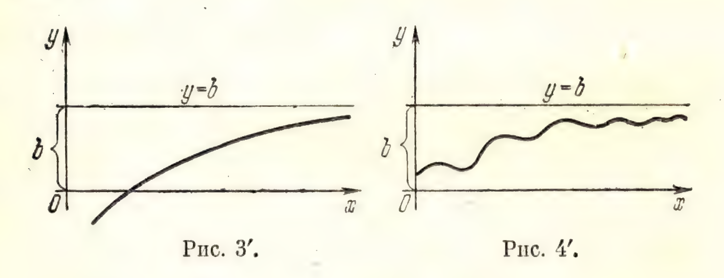


Рис. 3′. Рис. 4′.

В случае, когда a — число, а $b = +\infty$, возможны, очевидно, при $x \to a + 0$ только два типа, изображенных на рис. 6 и 7.

Предел постоянной величины. Вы, конечно, знаете, что предел постоянной величины, т. е. величины, принимающей одно-единственное значение, считается всегда существующим и равным этому единственному значению; такое соглашение иногда вызывает недоумение; ведь мы вначале говорим о пределе *переменной* вели-

чины; как же в таком случае может существовать предел у величины постоянной?

Однако такое соглашение, во-первых, логически обязательно для нас, если мы не хотим вступать в противоречие с нашим определением предела; в самом деле, если $y = b$ при любом значении величины x, то, какова бы ни была окрестность V точки b, мы имеем при любом x соотношение $y \in V$, откуда по определению предела мы обязаны заключить, что при любом a

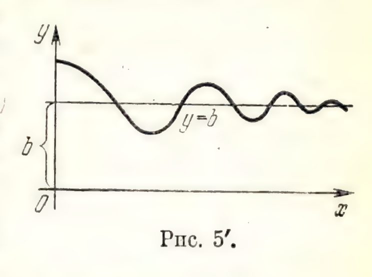


Рис. 5'.

$$\lim_{x \to a} y = b. \qquad (2)$$

Но такое соглашение, во-вторых, весьма целесообразно (в сущности, почти неизбежно) и практически;

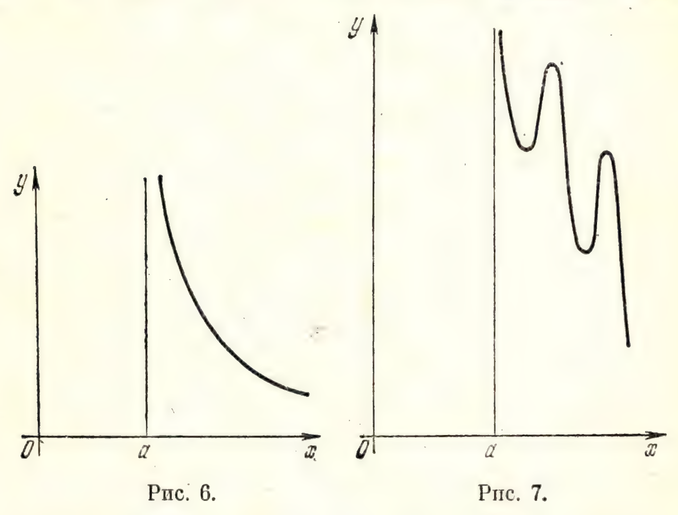


Рис. 6. Рис. 7.

в самом деле, допустим на минуту, что мы условились считать постоянную величину не имеющей предела,

и пусть для некоторой функции имеет место соотношение (2); тогда, как известно,

$$\lim_{x \to a} (-y) = -b.$$

Но $y + (-y) = 0$ при любом x; сумма величин y и $-y$, каждая из которых имеет предел, оказывается величиной постоянной и, значит, по нашему соглашению, не имеющей предела. Таким образом, формулируя известную теорему о пределе суммы двух переменных величин, мы вынуждены внести в нее оговорку: «если только эта сумма не есть постоянная величина»; подобные же оговорки, искусственность и обременительность которых ясны сами собой, нам пришлось бы внести почти во все теоремы теории пределов. От всего этого мы полностью освобождаемся, приписывая предел всякой постоянной величине. А так как при этом, как мы только что видели, мы не совершаем никакой погрешности против установленного нами определения понятия предела, то это соглашение всегда принимается при изложении основ математического анализа.

Бесконечно малые и бесконечно большие. Величина $y = f(x)$ называется *бесконечно малой* при $x \to a$, если $y \to 0$ при $x \to a$. Во многих изложениях анализа понятие бесконечно малой играет в теории пределов основную роль; оно определяется раньше, чем общее понятие предела, и, напротив, определение последнего редуцируется к понятию бесконечно малой величины. В противоположность этому некоторые современные ученые рекомендуют при изложении курса анализа обходиться вовсе без понятия бесконечно малой величины, полагая, что понятие это, будучи излишним, в то же время способно посеять, и фактически сеет, множество недоразумений.

По этому поводу следует заметить, что речь идет здесь, конечно, не о *понятии* величины, стремящейся к нулю (это понятие при любом построении анализа, а также и во всех его приложениях играет очень существенную роль), а лишь о *термине* «бесконечно малая величина». Это название действительно весьма неудачно и часто дает повод к недоразумениям; оно как бы

стремится охарактеризовать *размер* той величины, к которой оно относится,— как-то неудобно называть бесконечно малой величину, которая на известных стадиях своего изменения может быть и вовсе не малой; между тем на самом деле оно, конечно, призвано описывать только *характер изменения* данной величины, а отнюдь не ее размеры; эта терминологическая неудача является следствием исторического пережитка: бесконечно малые величины получили свое наименование в эпоху, когда смысл этого понятия был совершенно иным, чем мы его устанавливаем теперь.

Однако величины, стремящиеся к нулю, встречаются в анализе и его приложениях, как мы уже заметили, столь часто, что отказаться от специального краткого наименования для них было бы очень трудно. С другой стороны, заменять исторически прочно, веками утвердившееся наименование каким-либо другим означало бы такую терминологическую катастрофу, какая в математике всегда ощущается очень продолжительно и болезненно. Наконец, если, как это делаем сейчас мы, не принимать бесконечно малую величину за основание теории пределов, а понимать ее лишь как частный случай общего понятия величины, стремящейся к пределу, то опасность злокачественного смешения не так уж велика и не грозит столь тяжелыми последствиями.

Величина y называется *бесконечно большой*, если

$$\lim_{x \to a} |y| = +\infty.$$

Таким образом, бесконечно большая величина либо стремится к $+\infty$, либо стремится к $-\infty$, либо не стремится ни к какому пределу, принимая значения то положительные, то отрицательные, но по абсолютной величине безгранично возрастающие. Примером поведения последнего типа служит величина $y = (-1)^n n$ при $n \to +\infty$, причем n, возрастая, принимает только целые значения. О самом термине «бесконечно большая» можно повторить все то, что выше сказано о «бесконечно малой».

Нам нет надобности останавливаться здесь на хорошо известных вам теоремах о сумме, разности и произ-

ведении бесконечно малых величин. Напомним только два момента, недостаточное внимание к которым часто приводит к недоразумениям. 1) В теоремах о сумме и произведении бесконечно малых очень существенно требование, чтобы число слагаемых (или сомножителей) оставалось ограниченным; без этой оговорки обе теоремы становятся неверными, как легко убедиться на простых примерах. 2) Об *отношении* двух бесконечно малых нельзя сделать никакого общего утверждения; такое отношение может иметь любой характер изменения, в частности, может и не стремиться ни к какому пределу.

Прежде чем мы пойдем дальше, нам надо сделать еще несколько замечаний о символах $+\infty$ и $-\infty$, с которыми мы теперь уже достаточно часто встречались. Вы знаете, конечно, что эти символы не служат обозначением для каких-либо чисел. На вопрос о том, что означает, например, символ $+\infty$, правильнее и яснее всего будет ответить, что, взятый сам по себе, он лишен всякого смысла; смысл получает только выражение «окрестность $+\infty$», и под этим выражением разумеется совокупность всех вещественных чисел, превышающих некоторое (любое) число α; иметь для такой совокупности краткое наименование очень удобно, так как с совокупностями такого рода мы в анализе встречаемся на каждом шагу. Но вместе с выражением «окрестность $+\infty$» тотчас же получают смысл и такие выражения, как $\lim\limits_{x\to a} y = +\infty$ (где само a может также быть одним из символов $+\infty$ и $-\infty$); в самом деле, в данном нами определении предела $\lim\limits_{x\to a} y = b$ о буквах a и b упоминается исключительно в связи с их окрестностями; поэтому любая из этих букв может быть со смыслом заменена символом $+\infty$ (или $-\infty$) после того, как мы установили, какие множества мы называем окрестностями этого символа; при этом нет надобности приписывать какой-либо смысл самому этому символу.

Из всего этого ясно, с какой осторожностью необходимо относиться к употреблению символов $+\infty$ и $-\infty$; в частности, совершенно недопустимо применять к ним

какие-либо арифметические операции $\left(\frac{1}{\infty}=0 \text{ и т. п.}\right)$, как это принято в некоторых «упрощенных» курсах анализа. Равным образом лишены смысла и всякого рода равенства, в которых символ $+\infty$ или $-\infty$ выступает не в прямо указанной роли предела, как, например, распространенные в тригонометрии записи вида $\operatorname{tg}\frac{\pi}{2}=\pm\infty$ [1]) и т. п. Напротив, неравенствам вида $a<+\infty$, $b>-\infty$, $-\infty<c<+\infty$ придается определенный смысл; они последовательно означают: 1) что a есть либо число, либо символ $-\infty$; 2) что b есть либо число, либо символ $+\infty$; 3) что c есть число.

Наконец, необходимо точно условиться, будем ли мы, говоря о существовании предела функции $y=f(x)$, подразумевать при этом существование *числа*, служащего пределом этой величины, или будем допускать в качестве пределов также и символы $+\infty$, $-\infty$. Ясно, конечно, что в этом терминологическом вопросе мы формально можем принять любое из двух решений и что, следовательно, выбор должен быть принят на основе целесообразности. Обычно соглашаются понимать утверждение «y имеет предел» в смысле существования *числа*, служащего пределом величины y; если же имеется в виду широкое понимание этого утверждения, то делаются специальные оговорки. Так мы и будем поступать в дальнейшем.

Критерий Коши. Одной из важнейших задач теории пределов является, очевидно, следующая: *дана функция* $f(x)=y$; *определить, имеет ли она предел при* $x\to a$. При этом: 1) согласно только что принятому нами соглашению речь идет о существовании такого *числа* b, что $y\to b$ при $x\to a$; 2) напротив, a может означать либо число, либо один из символов $+\infty$, $-\infty$; 3) в этой задаче ставится вопрос только о *существовании* предела; отыскание же этого предела в данный момент нас совсем не должно занимать.

---

[1]) Правильная формулировка, конечно, такова:

$$\lim_{x\to\frac{\pi}{2}-0}\operatorname{tg}x=+\infty,\qquad \lim_{x\to\frac{\pi}{2}+0}\operatorname{tg}x=-\infty.$$

Следующий важный признак существования предела способен оказать весьма существенные услуги, в особенности в теоретических исследованиях.

Критерий Коши. *Для того чтобы функция $y = f(x)$ имела предел при $x \to a$, необходимо и достаточно выполнение следующего условия:* (A) *как бы мало ни было положительное число* ε, *существует такая окрестность* U *числа (или символа)* a, *что для любых двух чисел* x_1 *и* x_2 *окрестности* U, *отличных от* a,

$$|f(x_1)-f(x_2)| < \varepsilon.$$

Доказательство. 1) Пусть при $x \to a$ величина y стремится к числу b. По определению предела существует такая окрестность U числа a, что при $x \in U$ и $x \neq a$

$$b-\frac{\varepsilon}{2} < f(x) < b+\frac{\varepsilon}{2};$$

поэтому, если $x_1 \in U$ и $x_2 \in U$, причем $x_1 \neq a$ и $x_2 \neq a$, то $f(x_1)$ и $f(x_2)$ заключены между $b-\frac{\varepsilon}{2}$ и $b+\frac{\varepsilon}{2}$, а следовательно, отличаются друг от друга менее чем на ε. Это доказывает необходимость условия (A).

2) Допустим теперь, что условие (A) выполнено. Для каждого натурального числа n существует такая окрестность U_n числа a, что при $x_1 \in U_n$, $x_2 \in U_n$, причем $x_1 \neq a$ и $x_2 \neq a$, мы всякий раз имеем

$$|f(x_1)-f(x_2)| < \frac{1}{n}. \tag{3}$$

Мы имеем право предполагать, что $U_{n+1} \subset U_n$ ($n = 1, 2, \ldots$); в самом деле, если бы окрестность U_{n+1} не составляла части окрестности U_n, то мы просто заменили бы окрестность U_{n+1} общей частью U'_{n+1} окрестностей U_n и U_{n+1}[1]); очевидно, что U'_{n+1} есть окрестность, причем $U'_{n+1} \subset U_n$ и для любых $x_1 \in U'_{n+1}$, $x_2 \in U'_{n+1}$, если $x_1 \neq a$ и $x_2 \neq a$, мы имеем

$$|f(x_1)-f(x_2)| < \frac{1}{n+1},$$

---

[1]) Которая, очевидно, не пуста.

так что окрестность U'_{n+1} может во всех отношениях заменить окрестность U_{n+1}.

Так как функция $f(x)$ в окрестности U_n подчинена условию (3), то все множество M_n значений, принимаемых ею в этой окрестности, может быть заключено внутрь некоторого закрытого отрезка Δ_n длины $\frac{2}{n}$. Но $U_{n+1} \subset U_n$; следовательно, $M_{n+1} \subset M_n$, а потому отрезок Δ_{n+1} может быть взят целиком заключенным в отрезке Δ_n: $\Delta_{n+1} \subset \Delta_n$; а так как длина $\frac{2}{n}$ отрезка Δ_n стремится к нулю при $n \to \infty$, то последовательность отрезков Δ_n (1, 2, ...) является стягивающейся; в силу леммы 2 лекции I (стр. 32) существует единственное число b, принадлежащее всем отрезкам Δ_n.

Наконец, убедимся, что $\lim\limits_{x \to a} y = b$. С этой целью обозначим через V любую окрестность числа b; по определению этого числа $\Delta_n \subset V$ для всех достаточно больших n; поэтому, если $x \in U_n$ и $x \neq a$, то $f(x) \in M_n \subset \subset \Delta_n \subset V$, откуда по определению предела и следует, что $\lim\limits_{x \to a} y = b$.

Таким образом, критерий Коши доказан полностью. В частности, для того чтобы последовательность чисел

$$a_1, a_2, \ldots, a_n, \ldots$$

имела предел, необходимо и достаточно, чтобы при любом $\varepsilon > 0$ и всех достаточно больших n и m

$$|a_n - a_m| < \varepsilon.$$

Критерий Коши лишь в редких случаях получает конкретное применение в смысле доказательства существования предела у некоторой индивидуальной функции; с этой целью обычно применяются критерии более простые, но зато не являющиеся характеристическими, т. е. одновременно необходимыми и достаточными. Напротив, в общих теоретических исследованиях эта особенность критерия Коши делает его, как мы увидим дальше, почти незаменимым орудием.

Замечание об основных теоремах. Основные теоремы теории пределов, т. е. теоремы о пределе суммы,

разности, произведения и т. д., вам, конечно, хорошо известны; нам нет надобности здесь не только доказывать, но даже формулировать их. Однако одно замечание, связанное с этим циклом теорем, нам все же придется сделать; при этом, так как оно будет в равной мере относиться ко всем теоремам этого цикла, нам достаточно будет иллюстрировать его на какой-нибудь одной из них.

Когда мы говорим «предел суммы двух величин равен сумме их пределов» (а именно в этой краткой редакции теорема встречается в большинстве «упрощенных» курсов), то тем самым мы подразумеваем (хотя и не упоминаем этого), что все три величины (оба слагаемых и сумма) стремятся к пределам, и вопрос идет лишь о взаимоотношении этих пределов. На самом деле мы при этом предполагаем больше, чем нужно; достаточно предположить существование предела у каждого из слагаемых; тогда можно доказать, что и сумма должна иметь предел (равный сумме пределов слагаемых); доказательство это не требует никаких дополнительных рассуждений, автоматически вытекая в виде дополнительного результата при любом доказательстве теоремы о пределе суммы.

Напротив, из того, что сумма имеет предел, еще не вытекает существование предела у каждого из слагаемых; в самом деле, пусть величина y при $x \to a$ не стремится ни к какому пределу; очевидно, что то же самое будет иметь место для величины $1 - y$; между тем сумма этих двух величин, будучи постоянной ($y + 1 - y = 1$) имеет предел 1 при $x \to a$.

Таким образом, наиболее совершенная формулировка теоремы о пределе суммы должна гласить: *если каждая из данной конечной совокупности величин имеет предел при $x \to a$, то и сумма их имеет предел при этом условии, и предел суммы равен сумме пределов слагаемых*. Аналогичную формулировку получают и все другие теоремы этого цикла.

Частичные пределы; верхний предел и нижний предел. Мы теперь должны подробно заняться изучением таких функций величины x, которые при $x \to a$ не имеют предела; так называемые «упрощенные» курсы ос-

тавляют обычно этот вопрос совершенно без рассмотрения.

Условимся называть число b *частичным пределом* функции $y = f(x)$ при $x \to a$, если, каковы бы ни были окрестность U числа a и окрестность V числа b, найдется такое $x \in U$, отличное от a, что $f(x) \in V$. Это простое определение без всяких изменений распространяется и на тот случай, когда b есть один из символов $+\infty$, $-\infty$. Таким образом, частичный предел b есть такое число, что как угодно близко к a найдутся значения величины x, при которых величина y как угодно близка к b [в случае, когда a или b есть $+\infty$, мы должны, конечно, вместо чисел «как угодно близких к a (или b)» говорить о числах «как угодно больших»; аналогично и в случае, когда a или b есть $-\infty$].

Аналогично определяется частичный предел последовательности $a_1, a_2, \ldots, a_n, \ldots$ при $n \to \infty$.

Если $\lim\limits_{x \to a} y$ существует, то он, как мы непосредственно видим из определения понятия предела, есть вместе с тем частичный предел (и притом единственный) величины y, независимо от того, является ли этот предел числом или одним из символов $+\infty$, $-\infty$; но если этого предела нет, то величина y необходимо имеет по меньшей мере два различных частичных предела; таким образом, любая функция $y = f(x)$ имеет при $x \to a$ по меньшей мере один частичный предел; для того чтобы он был единственным, необходимо и достаточно существование $\lim\limits_{x \to a} y$.

Чтобы убедиться в справедливости всех этих утверждений, мы, очевидно, должны доказать два предложения: 1) *любая функция $y = f(x)$ имеет при $x \to a$ по меньшей мере один частичный предел и* 2) *если b есть единственный частичный предел функции $y = f(x)$ при $x \to a$, то*

$$\lim_{x \to a} y = b.$$

При этом во всех случаях все пределы и частичные пределы могут быть как числами, так и символами $+\infty$ и $-\infty$.

Доказательство утверждения 1). Если один из символов $+\infty$ или $-\infty$ служит частичным пределом величины y при $x \to a$, то утверждение доказано; поэтому мы допустим, что это не так, и покажем, что в таком случае существует число b, являющееся частичным пределом величины y при $x \to a$.

Но если ни $+\infty$, ни $-\infty$ не служит частичным пределом функции $f(x)$ при $x \to a$, то, очевидно, существует такая пара чисел α, β ($\alpha < \beta$), что для всех $x \neq a$, достаточно близких к a,

$$\alpha \leqslant f(x) \leqslant \beta.$$

В частности, значит, *в любой окрестности числа a найдется такое число $x \neq a$, что $f(x)$ принадлежит отрезку* $[\alpha, \beta]$, который мы для краткости обозначим через Δ_1; условимся обозначать через A только что сформулированное *свойство* отрезка Δ_1; если мы разделим этот отрезок пополам, то по меньшей мере одна из его половин обладает свойством A (ибо, если бы, например, окрестность U_1 числа a не содержала ни одного $x \neq a$, для которого $f(x)$ принадлежит левой половине Δ_1, а окрестность U_2 — ни одного такого, для которого $f(x)$ принадлежит правой половине этого отрезка, то окрестность U, служащая общей частью U_1 и U_2, очевидно, не могла бы содержать ни одного $x \neq a$, для которого $f(x) \in \Delta_1$, т. е. отрезок Δ_1 не обладал бы свойством A). Обозначим через Δ_2 половину отрезка Δ_1, обладающую свойством A, и поступим с ней так, как мы поступили с отрезком Δ_1, т. е. разделим ее пополам и возьмем ее половину Δ_3, обладающую свойством A. Продолжая этот процесс безгранично, мы, очевидно, получаем стягивающуюся последовательность отрезков $\Delta_1, \Delta_2, \Delta_3, \ldots$ $\ldots, \Delta_n, \ldots$ Пусть b — существующее в силу леммы о стягивающейся последовательности отрезков единственное общее число всех этих отрезков, и пусть U и V означают соответственно любые окрестности чисел a и b; по определению числа b найдется отрезок $\Delta_n \subset V$; но каждый отрезок Δ_n обладает свойством A, следовательно, существует такое отличное от a число $x \in U$, что $f(x) \in \Delta_n \subset V$. Но это и значит, что b есть частичный предел величины y при $x \to a$.

Доказательство утверждения 2). В силу утверждения 1) величина y имеет при $x \to a$ по меньшей мере один частичный предел b; если соотношение $\lim_{x \to a} y = b$ неверно, то существует окрестность V числа (или символа) b, обладающая тем свойством, что в любой окрестности U числа (или символа) a найдется такое число $x \neq a$, для которого $f(x)$ лежит вне V. Очевидно поэтому, что должна иметь место одна из следующих двух ситуаций: либо один из символов $+\infty$, $-\infty$, и притом *отличный от* b, является частичным пределом величины y при $x \to a$, либо существует отрезок $[\alpha, \beta]$, *лежащий вне* V и такой, что любая окрестность U числа (или символа) a содержит число $x \neq a$, для которого $f(x) \in [\alpha, \beta]$. В первом случае утверждение 2) доказано. Во втором случае мы в точности тем же путем, как при доказательстве утверждения 1), убедимся, что величина y имеет при $x \to a$ частичный предел b', принадлежащий отрезку $[\alpha, \beta]$ и, следовательно, отличный от b; таким образом, и в этом случае утверждение 1) доказано.

Итак, мы установили, что всякая величина, которая в данном процессе не имеет предела, должна обладать по меньшей мере двумя частичными пределами. Вообще говоря, ничего большего утверждать нельзя; так, величина $y = (-1)^n$ (где n — натуральное число) имеет при $n \to \infty$, очевидно, в точности два частичных предела 1 и -1. С другой стороны, существуют величины, которые в определенных процессах изменения могут иметь и бесконечное множество частичных пределов; так, функция $y = \sin \frac{1}{x}$ имеет при $x \to 0$ в качестве частичных пределов все числа отрезка $[-1, +1]$; в самом деле, при $x \to 0$ величина $\frac{1}{x}$, возрастая непрерывно и безгранично, имеет синус, бесчисленное множество раз непрерывно изменяющийся от -1 до $+1$ и обратно; каково бы ни было число β отрезка $[-1, +1]$, мы можем, следовательно, найти сколь угодно малое (т. е. принадлежащее как угодно малой окрестности числа 0) число α, для которого $\sin \frac{1}{\alpha} = \beta$.

Несмотря на то, что различные величины, как мы только что видели, в смысле множества имеющихся у них частичных пределов являют собой весьма различную картину, все же в строении множества частичных пределов имеются черты, общие для всех переменных величин, и черты эти как раз для анализа имеют существенную важность. На некоторых из них мы должны будем теперь остановиться.

Свойство 1. *Если число (или символ) b не является частичным пределом величины $y=f(x)$ при $x \to a$, то существует окрестность V числа (или символа) b, не содержащая внутри себя ни одного такого частичного предела.*

В самом деле, если b не есть частичный предел величины y при $x \to a$, то существуют такая окрестность V числа (или символа) b и такая окрестность U числа (или символа) a, что y не может принадлежать открытому отрезку V при $x \in U$, если $x \neq a$; но так как отрезок V является окрестностью любого своего внутреннего числа, то в точности то же самое имеет место и для любого такого внутреннего числа; ни одно из этих чисел не может, следовательно, быть частичным пределом величины y при $x \to a$, чем и доказывается свойство 1.

Прежде чем перейти ко второму свойству, заметим следующее.

Если среди частичных пределов данной величины y при $x \to a$ встречается символ $+\infty$, то мы, естественно, принимаем этот частичный предел за наибольший; подобным же образом символ $-\infty$, если он служит частичным пределом нашей величины, мы считаем наименьшим из ее частичных пределов. Условившись в этом, мы можем теперь формулировать свойство 2.

Свойство 2. *Среди частичных пределов переменной величины $y = f(x)$ при $x \to a$ всегда имеются наибольший и наименьший.*

Вы понимаете, конечно, что это утверждение далеко не самоочевидно; ведь нет же, например, ни наибольшего, ни наименьшего числа в открытом отрезке $(0, 1)$, и вообще далеко не во всяком множестве чисел имеются наибольшее и наименьшее числа (или символы); смысл нашего утверждения в том и заключается, что множест-

вом частичных пределов данной переменной величины может быть отнюдь не любое множество (что, конечно, видно из свойства 1); оно должно обладать некоторыми специфическими чертами и, в частности, содержать наибольшее и наименьшее числа (или символы); докажем это.

Прежде всего, будем говорить только о *наибольшем* частичном пределе, так как для наименьшего (это заранее очевидно) все рассуждения будут протекать совершенно аналогично.

Мы можем допустить, что символ $+\infty$ не является частичным пределом данной переменной величины (иначе он, как уже замечено, был бы наибольшим частичным пределом, и свойство 2 было бы доказано). В силу свойства 1 отсюда следует, что частичных пределов нет и в некоторой окрестности символа $+\infty$, другими словами, все частичные пределы меньше некоторого числа. Если бы единственным частичным пределом нашей величины был символ $-\infty$, то он был бы и наибольшим частичным пределом, и свойство 2 было бы доказано. Поэтому мы можем предположить, что среди частичных пределов данной величины имеется некоторое число b.

Разобьем теперь совокупность всех вещественных чисел на два класса A и B по следующему принципу: $\gamma \in A$, если вправо от γ имеется хотя бы один частичный предел данной величины, и $\gamma \in B$ — в противном случае. Вы сами легко убедитесь, что это разбиение есть сечение. Пусть α — его рубеж; покажем, что α есть наибольший частичный предел данной величины. В самом деле, прежде всего, α есть частичный предел, так как в противном случае в силу свойства 1 некоторая окрестность (α_1, α_2) числа α не содержала бы ни одного частичного предела; но в силу неравенства $\alpha_1 < \alpha$ мы имеем $\alpha_1 \in A$, так что вправо от α_1 должны быть частичные пределы; так как их нет в отрезке (α_1, α_2), то они должны быть больше или равны α_2, что невозможно, так как любое число α_3, для которого $\alpha < \alpha_3 < \alpha_2$, принадлежит B. Далее, вправо от α частичных пределов быть не может, так как если $\beta > \alpha$ есть частичный предел, то для любого γ, заключенного между α и β, $a < \gamma < \beta$, мы имели бы $\gamma \in B$ в силу $\gamma > \alpha$ и в то же

время $\gamma \in A$ в силу $\beta > \gamma$. Таким образом, число α действительно есть наибольший частичный предел данной величины, и свойство 2 доказано.

Итак, любая переменная величина при $x \to a$ имеет наибольший и наименьший частичные пределы. Эти два частичных предела имеют важное значение; их называют соответственно *верхним* и *нижним* пределами величины $y = f(x)$ при $x \to a$ и обозначают через

$$\overline{\lim_{x \to a}} f(x) \text{ и } \underline{\lim_{x \to a}} f(x)$$

или

$$\limsup_{x \to a} f(x) \text{ и } \liminf_{x \to a} f(x).$$

Очевидно, всегда

$$\limsup_{x \to a} f(x) \geqslant \liminf_{x \to a} f(x);$$

в остальном эти два числа могут принимать какие угодно значения, причем каждое из них или даже оба могут быть и символами $+\infty$ или $-\infty$. Величина y называется *ограниченной* при $x \to a$, если верхний и нижний пределы ее при $x \to a$ являются числами, и *неограниченной*, если хотя бы один из них есть символ $+\infty$ или $-\infty$; очевидно, ограниченность величины y при $x \to a$ означает существование таких чисел α, β ($\alpha < \beta$) и такой окрестности U числа (или символа) a, что при любом $x \in U$, отличном от a,

$$\alpha < f(x) < \beta.$$

Для существования $\lim\limits_{x \to a} y$ (в смысле числа или символа) необходимо и достаточно соотношение

$$\overline{\lim_{x \to a}}\, y = \underline{\lim_{x \to a}}\, y,$$

ибо совпадение верхнего предела с нижним равносильно единственности частичного предела.

Числа, заключенные между $\underline{\lim\limits_{x \to a}}\, y$ и $\overline{\lim\limits_{x \to a}}\, y$, могут как быть, так и не быть частичными пределами величины y при $x \to a$; выше мы видели примеры двух крайних

случаев, когда 1) ни одно из этих чисел не служит частичным пределом и 2) напротив, каждое из них является частичным пределом. В общем случае некоторые из них будут частичными пределами, а другие — нет.

Иногда верхний и нижний пределы определяются иначе, и эти иные определения полезно знать как для расширения и конкретизации самого представления об этих понятиях, так и для приложений. *Число (или символ) b называют верхним пределом величины $y = f(x)$ при $x \to a$, если, каковы бы ни были числа (или символы) α, β $(\alpha < b < \beta)$, найдется сколь угодно малая окрестность U числа (или символа) a, содержащая числа $x \neq a$, для которых $\alpha \leqslant y \leqslant \beta$, но не содержащая таких чисел $x \neq a$, для которых $y > \beta$.*

Чтобы показать, что новое определение равносильно прежнему, заметим прежде всего, что если число или символ b есть верхний предел величины y по новому определению, то b, очевидно, является частичным пределом этой величины; если бы существовал частичный предел $b' > b$, то мы могли бы найти числа α, β, α', β' так, что

$$\alpha < b < \beta < \alpha' < b' < \beta';$$

при этом, так как b' есть частичный предел, любая окрестность U числа (или символа) a должна содержать число $x \neq a$, для которого $\alpha' \leqslant y \leqslant \beta'$, откуда $y > \beta$; но это, очевидно, противоречит тому, что b есть верхний предел согласно новому определению.

Пусть теперь, обратно, b есть наибольший частичный предел величины y при $x \to a$; если $\alpha < b < \beta$, то любая окрестность U числа (или символа) a содержит число $x \neq a$, для которого $\alpha < y < \beta$ (по определению частичного предела). Чтобы убедиться, что b есть верхний предел величины y в смысле нового определения, остается поэтому только показать, что при надлежащем выборе окрестности U мы для любого $x \in U$ при $x \neq a$ будем иметь $y \leqslant \beta$.

Если $b = +\infty$, то доказывать нечего; если нет, то $+\infty$ не может быть частичным пределом (ибо b есть *наибольший* частичный предел). Значит, существуют такое число y_0 и такая окрестность U_0 числа (или сим-

вола) a, что $y < y_0$ для всех $x \neq a$, принадлежащих U_0. Если $y_0 \leqslant \beta$, то все требуемое доказано. Если $y_0 > \beta$, то каждое число λ отрезка $[\beta, y_0]$, не будучи частичным пределом, имеет окрестность V_λ, которой соответствует такая окрестность U_λ числа (или символа) a, что для любого $x \in U_\lambda$ при $x \neq a$ значение $f(x) = y$ лежит вне V_λ. Открытые отрезки V_λ покрывают отрезок $[\beta, y_0]$. Применяя лемму Гейне — Бореля (лемма 3 лекции I), мы находим конечную группу окрестностей $V_{\lambda_1}, V_{\lambda_2}, \ldots$ $\ldots, V_{\lambda_r}$, покрывающую отрезок $[\beta, y_0]$. Пусть тогда U есть общая часть окрестностей $U_0, U_{\lambda_1}, U_{\lambda_2}, \ldots, U_{\lambda_r}$; очевидно, что U есть окрестность числа (или символа) a и что если $x \in U$ и $x \neq a$, то соответствующее значение y не может принадлежать ни одному из отрезков $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_r}$; другими словами, $y < \beta$ для любого $x \in U$, чем наше утверждение и доказано.

Само собой разумеется, что аналогичное новое определение получает и нижний предел и что равносильность старого и нового определений здесь доказывается совершенно аналогичным образом.

Если желательно бросить более конкретный взгляд на всю эту абстрактную схему, то нашим глазам открывается картина, которую без всякой претензии на формальную строгость можно попытаться описать следующими основными штрихами. Переменная величина, находясь в некотором процессе изменения, приближается при этом то к одному числу, то к другому, то к третьему и т. д. Мы называем число b частичным пределом данной величины, если она на сколь угодно далеких (поздних) стадиях процесса все вновь и вновь как угодно близко подходит к числу b (причем вполне возможно, что в промежутке между двумя такими приближениями она уклоняется от b весьма далеко; важно лишь, чтобы рано или поздно она снова подошла к b на сколь угодно малое расстояние). Если b есть *единственный* такой центр притяжения, то понятно, что наша величина y не только становится, но в конце концов и *остается* как угодно близкой к b, и мы имеем $\lim\limits_{x \to a} y = b$. Вообще же говоря, y не стремится ни к какому пределу, но, сколь бы далеко мы ни следили за процессом его

изменения, продолжает, подобно незатухающему маятнику, колебаться в известных границах; эти границы (среди которых могут встречаться $-\infty$ и $+\infty$) и представляют то, что мы называли верхним и нижним пределами данной величины.

Пределы функций нескольких переменных. Нам остается еще рассмотреть несколько несложных вопросов теории пределов, связанных с функциями нескольких переменных; мы будем при этом, чтобы упростить изложение, говорить лишь о функциях *двух* независимых переменных; все сказанное нами при этом может быть, с соответствующими очевидными изменениями, перенесено и на случай функций любого числа переменных.

Каждая пара значений независимых переменных x и y изображается точкой координатной плоскости xy, и мы в дальнейшем под именем *точки* (a, b) будем для краткости понимать пару значений $x = a$, $y = b$ этих переменных. Под окрестностью точки (a, b) мы будем понимать любую область плоскости xy [1]), содержащую эту точку внутри себя; эта область может быть кругом, прямоугольником или иметь более сложную форму. Термин «окрестность точки (a, b)» мы считаем сохраняющим свой смысл и в том случае, если a или b или обе буквы означают какой-либо из символов $+\infty$, $-\infty$; так, если a — число, а b — символ $+\infty$, то окрестностью точки (a, b) мы назовем любую область вида $\alpha<x<\beta$, $y>\gamma$, где $\alpha<a<\beta$, а число γ произвольно (рис. 8); если a есть $-\infty$, а b по-прежнему $+\infty$, то окрестностью точки (a, b) служит любая область вида $x<\alpha$, $y>\beta$, где α и β — произвольные числа (рис. 9), и т. д.

Мы говорим, что функция $z = f(x, y)$ при $x \to a$, $y \to b$ имеет своим пределом число (или символ) c, если для любой окрестности V числа (или символа) c суще-

---

[1]) Областью на плоскости называют множество точек плоскости, обладающее следующим свойством: если точка принадлежит этому множеству, то найдется круг с центром в этой точке, целиком содержащийся в том же множестве. [Обычно множество, обладающее указанным свойством, называют просто *открытым* множеством, *область* же — это открытое связное множество.— *Прим. ред.*]

ствует такая окрестность U точки (a, b), что в любой точке $(x, y) \in U$ при $(x, y) \neq (a, b)$ мы имеем $z \in V$. Записывается это положение вещей следующим образом:

$$\lim_{\substack{x \to a \\ y \to b}} z = c \quad \text{или} \quad \underset{\substack{x \to a \\ y \to b}}{z \to c}. \qquad (4)$$

При этом определении вся теория пределов, построенная для функций одной независимой переменной,

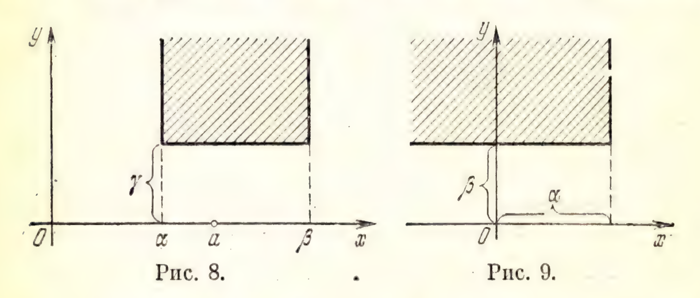


Рис. 8. Рис. 9.

легко и естественно переносится на двумерный случай; в частности, понятия частичного, верхнего и нижнего пределов сохраняют свои определения и свои свойства; точно так же в силе остается признак Коши и его доказательство.

Что касается обоснования, то оно происходит в точной аналогии с одномерным случаем; единственное отличие состоит в том, что теперь нам приходится пользоваться двумерными вспомогательными теоремами, например леммой о стягивающейся последовательности прямоугольников, двумерной леммой Гейне — Бореля и т. п.; все эти вспомогательные предложения доказываются так же просто и тем же путем, как и соответствующие одномерные, так что нам нет надобности здесь на них останавливаться.

От двойного предела (4) необходимо строго отличать *повторные* пределы

$$\lim_{y \to b}\left(\lim_{x \to a} z\right) \quad \text{и} \quad \lim_{x \to a}\left(\lim_{y \to b} z\right); \qquad (5)$$

здесь вместо одного двумерного мы имеем два последо-

вательных одномерных предельных перехода; геометрическая картина, связанная с окрестностями точки (a, b) в плоскости xy, здесь полностью исчезает; так, для получения $\lim\limits_{y\to b}\left(\lim\limits_{x\to a} z\right)$ мы должны сначала придать y какое-либо постоянное значение и, сделав таким образом величину z функцией *одной* переменной x, искать ее предел при $x \to a$; этот предел может существовать для одних закрепленных значений величины y и отсутствовать для других ее значений; если он существует для всех значений $y \neq b$, лежащих в некоторой окрестности V точки (или символа) b, то в этой окрестности $\lim\limits_{x\to a} f(x, y)$ есть функция одной переменной y, и мы можем искать ее предел при $y \to b$; этот последний, если он существует, и есть $\lim\limits_{y\to b}\left(\lim\limits_{x\to a} z\right)$.

Может случиться, что оба повторных предела (5) существуют, в то время как двойной предел (4) отсутствует. Рассмотрим, например, поведение функции

$$z = \frac{2xy}{x^2 + y^2}$$

в соседстве начала координат (то, что эта функция в самом начале координат не определена, не имеет здесь никакого значения, ибо в определении как двойного, так и повторных пределов величина $f(a, b)$ вовсе не участвует; если угодно, можно приписать функции z при $x = y = 0$ любое значение). При постоянном $x \neq 0$ будет $\lim\limits_{y\to 0} z = 0$; при постоянном $y \neq 0$ будет $\lim\limits_{x\to 0} z = 0$; следовательно,

$$\lim_{x\to 0}\left(\lim_{y\to 0} z\right) = \lim_{y\to 0}\left(\lim_{x\to 0} z\right) = 0.$$

С другой стороны, любая окрестность точки $(0, 0)$ содержит как точки, в которых $x = 0$, $y \neq 0$, так и точки, в которых $x = y \neq 0$; так как в точках первого типа $z = 0$, а в точках второго типа $z = 1$, то оба числа 0 и 1 являются частичными пределами функции z при $x \to 0$, $y \to 0$; наличие же по меньшей мере двух различных пределов показывает, что $\lim\limits_{\substack{x\to 0\\ y\to 0}} z$ не существует.

Может иметь место и обратное положение вещей, когда двойной предел (4) существует, в то время как нет ни одного из повторных пределов (5); рассмотрим функцию

$$x = \begin{cases} (x^2 + y^2)\sin\dfrac{1}{xy}, & \text{если } xy \neq 0, \\ 0, & \text{если } xy = 0, \end{cases}$$

в соседстве начала координат; так как $|z| < \varepsilon^2$ для всех точек круга $x^2 + y^2 < \varepsilon^2$, то $\lim\limits_{\substack{x\to 0\\ y\to 0}} z = 0$; с другой стороны, при постоянном $x \neq 0$ и при $y \to 0$ (а также при постоянном $y \neq 0$ и при $x \to 0$) $\sin\dfrac{1}{xy}$, а следовательно и функция z, не стремится ни к какому пределу; таким образом, при $y \neq 0$ предел $\lim\limits_{x\to 0} z$ и при $x \neq 0$ предел $\lim\limits_{y\to 0} z$ отсутствует, тем более, конечно, не может быть и речи о существовании какого-либо из повторных пределов (5).

Можно, однако, показать, что в случае, когда все три предела (4) и (5) существуют, они необходимо совпадают между собой. В самом деле, пусть, например,

$$\lim_{x\to a}\left(\lim_{y\to b} z\right) = c; \tag{6}$$

пусть V — любая окрестность числа (или символа) c и U — любая окрестность точки (a, b). В силу (6) существует такая окрестность A числа (или символа) a, что для всех $x \in A$ при $x \neq a$

$$\lim_{y\to b} z \in V;$$

эта окрестность A есть некоторый отрезок оси Ox; очевидно, мы можем выбрать на этом отрезке число x_0 столь близким к a, чтобы точка (x_0, b) принадлежала окрестности U (рис. 10). Закрепив это значение x_0, мы в силу $x_0 \in A$ имеем

$$c_{x_0} = \lim_{y\to b} z \in V;$$

но согласно определению понятия предела это означа-

ет, что для любой окрестности V' числа (или символа) c_{x_0} можно указать такую окрестность B_{x_0} числа (или символа) b, что $z \in V'$ при $x = x_0$ и любом $y \neq b$, принадлежащем B_{x_0}; но очевидно, ничто не препятствует нам выбрать окрестность V' столь малой, чтобы $V' \subset V$ (ибо $c_{x_0} \in V$); с другой стороны, так как точка $(x_0, b) \in U$, то мы можем выбрать число $y_0 \in B_{x_0}$ столь близким к b, чтобы и точка (x_0, y_0) принадлежала окрестности U (рис. 10).

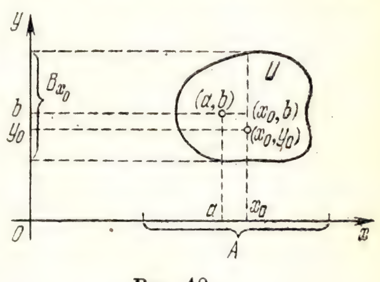


Рис. 10.

В силу $y_0 \in B_{x_0}$ при этом $z \in V' \subset V$ в точке (x_0, y_0). Вспомним теперь, что U — *любая* окрестность точки (a, b) и V — *любая* окрестность числа (или символа) c; то, что мы нашли точку $(x_0, y_0) \in U$, для которой $z \in V$, показывает поэтому, что c есть частичный предел величины z при $x \to a$, $y \to b$. Если, как мы допустили, $\lim\limits_{\substack{x \to a \\ y \to b}} z$ существует, то он является единственным частичным пределом и поэтому должен равняться c, что и требовалось доказать.

ЛЕКЦИЯ III

ФУНКЦИИ

Что такое функция.— Область определения функции.— Непрерывность.— Ограниченные функции.— Основные свойства непрерывных функций.— Непрерывность элементарных функций.— Колебание функции в точке.— Точки разрыва.— Монотонные функции.— Функции с ограниченной вариацией.

Что такое функция. То определение понятия функциональной зависимости, которое мы напомнили и в некоторых отношениях исследовали в самом начале первой лекции, родилось и окончательно восторжествовало в тяжелой борьбе; пережитки и рецидивы этой борьбы далеко не ликвидированы еще и в настоящее время, хотя, кажется, научная актуальность приведенного нами определения никем уже не оспаривается[1]). Основным лозунгом этой борьбы было и остается преодоление того засилия аналитического аппарата, которое с 18-го столетия тяготеет над идеей функциональной зависимости. Это засилие, это превращение аналитического выражения из удобного орудия исследования в деспотического властителя идеи функции, диктующего ей свои законы, сколько-нибудь полно изжито только в пределах самой математической науки; в прикладных науках и в преподавании (даже в высшей технической школе) оно еще в весьма значительной степени сохранило свое влияние, и почти каждый инженер, формально признавая научное определение функциональной зависимости (в котором ни единым словом не упоминается об аналитическом выражении),

[1]) Если не считать интуиционистской школы; впрочем, ее концепция функциональной зависимости во всяком случае не знаменует собой возврата к старым представлениям.

все же представляет себе функцию прежде всего как формулу, как аналитическое выражение и вне этой связи, как правило, вообще не умеет мыслить о функциональной зависимости; этим он существенно отличается от математика, привыкшего принимать всерьез определения своих понятий и потому при слове «функция» всегда мыслящего о том соответствии между двумя множествами, о котором говорится в определении этого понятия, никак не ассоциируя его с каким-либо аналитическим аппаратом. Вы поймете, что ввиду такого положения вещей нам придется несколько остановится на идее функциональной зависимости и некоторых связанных с нею понятиях и представлениях.

Будет, может быть, лучше всего, если мы, отказавшись от проведения полного и систематического анализа (на который у нас и не хватило бы времени), просто остановимся на нескольких типичных острых конфликтах, возникающих здесь между пониманием функциональной зависимости, с одной стороны, математиком, а с другой — человеком, воспитанным на устаревших традициях; конфликты этого рода мы, профессора университетов, ежегодно имеем с нашими студентами-первокурсниками, приносящими из воспитавшей их средней школы навыки и традиции минувших столетий.

Пусть речь идет об аналитическом выражении функции, графическое изображение которой мы имеем на рис. 11. Один взгляд на этот простой график убеждает нас, что мысль о реальной величине y, изменения которой происходят согласно указываемому им закону, не содержит в себе ничего неприемлемого; и в этом утверждении, пожалуй, инженер или физик будет чувствовать себя даже увереннее, чем математик.

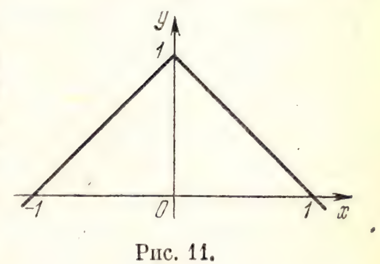


Рис. 11.

Но для поставленной задачи — выразить данную функцию аналитически — математик сразу напишет

единственное приемлемое, с его точки зрения, решение:

$$y = \begin{cases} 1 - x & (x \geqslant 0), \\ 1 + x & (x \leqslant 0). \end{cases} \qquad (1)$$

Вот тут и начинается конфликт. Инженер (условимся для краткости называть этим именем представителя устаревших традиций, присутствующие среди вас товарищи инженеры не осудят нас за это) немедленно возражает, заявляя математику, что тот «написал не одну, а две функции». На это математик отвечает: а) что «написать функцию» вообще нельзя, а можно писать только аналитические выражения; б) что он, математик, действительно написал два аналитических выражения, точно указав при этом, при каких значениях независимой переменной x следует употреблять то или другое из этих выражений для вычисления соответствующего значения функции y; в) что задание (1) в силу принятого определения функциональной зависимости определяет в точности *одну* функцию, ибо на основании этого задания каждому значению величины x соответствует только одно значение величины y; г) что, наконец, аналитическая запись (1) в точности воспроизводит зависимость, графически заданную на рис. 11, и тем самым полностью решает поставленную задачу. Математик мог бы еще добавить (но не делает этого из педагогических соображений), что в случае настойчивых требований он сумел бы представить функцию рис. 11 и одной формулой

$$y = 1 - |x|, \qquad (2)$$

годной для всех x, и не делает этого лишь потому, что считает задание (1) практически удобнее задания (2), в то время как с принципиальной стороны эти два задания в его глазах совершенно равноправны. Он идет еще дальше и настаивает, ссылаясь на определение функциональной зависимости, на том, что, например,

$$y = \begin{cases} x & \text{при } x < 0, \\ \frac{1}{2} & \text{при } x = 0, \\ 1 + x & \text{при } x > 0 \end{cases}$$

есть полноценное задание одной-единственной функции (рис. 12).

Мы не знаем, что может возразить на все это инженер, но на основе личного опыта тем не менее беремся сделать прогноз: если он и ничего не возразит, он все же уйдет неудовлетворенным; привитые годами навыки не изжить, конечно, одной краткой дискуссией.

В другой раз математик расскажет определение так называемой «функции Дирихле»:

$$y = \begin{cases} 1, & \text{если } x \text{ — рациональное число,} \\ 0, & \text{если } x \text{ — иррациональное число.} \end{cases}$$

Инженер в недоумении спросит: «Какая же это функция? Ее нельзя ни записать формулой, ни изобразить графиком!» На это математик возразит: а) что в определении функциональной зависимости ни словом не упоминается ни об аналитическом выражении, ни о геометрическом изображении функции и что поэтому полноценность задания функциональной зависимости нисколько не зависит от того, может ли определяемая этим заданием функция быть выражена аналитически или изображена геометрически; б) что сформулированное им определение функции Дирихле относит каждому значению величины x единственное значение величины y и потому является безукоризненно полноценным; в) что, наконец, геометрическое изображение функции Дирихле действительно затруднительно, в то время как выразить ее формулой весьма просто: для этого достаточно обозначить ее через $f(x)$.

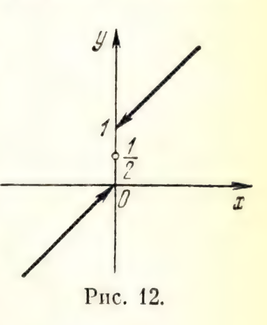


Рис. 12.

Инженер, вначале слушавший спокойно, при последнем замечании испытывает взрыв искреннего возмущения, в результате которого возникает следующий диалог:

И.— Да разве же это — формула?

М.— Что вы называете формулой?

И.— Ну, какое-нибудь аналитическое выражение $y = x^2 - x^3$ или $y = \sin x$, но какая же формула $y = f(x)$?

М.— Очень хорошо. Значит, запись $y = \sin x$ для известной функции, называемой синусом, по-вашему, является аналитическим выражением, а запись $y = f(x)$ для функции Дирихле таковым не является? Какое же принципиальное различие заставляет вас столь различно оценивать символы sin и $f(\ )$?

И.— Но ведь каждый грамотный человек знает, что означает формула $y = \sin x$, а обозначение $y = f(x)$ для того, что вы называете функцией Дирихле, сейчас вами придумано и никому не известно!

М.— Вот мы с вами, кажется, и договорились; указанное вами различие (которого я не отрицаю), как вы сами понимаете, имеет не принципиальный, а исторический характер. Ведь был такой момент, когда кто-то в первый раз предложил обозначить через $\sin x$ функцию, для которой до тех пор никакого общепринятого обозначения не было; этот момент для функции $y = \sin x$ был тем же, чем настоящий момент является для функции Дирихле $y = f(x)$; скажете ли вы, что до этого момента функция $y = \sin x$ была аналитически невыразимой, а после него, особенно когда предложенное обозначение привилось и стало общепринятым, сделалась аналитически выразимой? И если предложенное мною сегодня обозначение $y = f(x)$ для функции Дирихле через несколько лет будет принято всем научным миром, скажете ли вы тогда, что оно стало формулой, аналитическим выражением, а сама функция Дирихле — аналитически выразимой? Вам ясно, конечно, что такое апеллирующее к научной моде понимание аналитической выразимости ни малейшего математического значения не имеет; если же от этого понимания отказаться, то вы не можете не признать, что символ $f(\ )$ принципиально столь же полноценен, как и символ sin, и что, стало быть, функция Дирихле в такой же мере и в таком же смысле аналитически выразима, как синус и косинус; вообще дискуссия об аналитической выразимости той или другой функции во всех случаях должна быть признана бессодержа-

тельной, так как мы можем обозначить эту функцию любым символом и с полным основанием считать это обозначение ее аналитическим выражением. Наконец, могу вам сообщить, что мы умеем выражать функцию Дирихле и с помощью привычных вам символов; но мы почти никогда не пользуемся этим выражением, потому что оно сложно и не дает возможности узнать что-нибудь существенное о свойствах этой функции, в то время как данное нами бесформульное определение показывает все эти свойства, как на ладони. Мы вообще не любим фетишизировать аналитических выражений; мы охотно пользуемся ими в тех случаях, когда они помогают нам исследовать данную функциональную зависимость; и мы без всякого сожаления отбрасываем их, как негодное орудие, там, где исследование проще вести без них, как это имеет место, например, в только что рассмотренном нами случае.

Мы должны признать дискуссию на этом законченной, так как мы снова не видим, что бы мог возразить наш инженер; с другой стороны, наш математик с такой четкостью охарактеризовал отношение современной математической науки к связи между функцией и аналитическим выражением, что выводы ясны сами собой и не нуждаются ни в каких комментариях.

Область определения функции. Коснемся еще только одного вопроса, тесно связанного с предыдущим. Для того чтобы мы могли считать функцию $y = f(x)$ заданной, вовсе не всегда необходимо определить ее значения для всех без исключения значений величины x; очень часто случается, что сколько-нибудь актуальным бывает только тот или другой участок всей совокупности этих значений, так что определять функцию $f(x)$ вне этого участка представляется бесцельным, а иногда и нелепым; при этом поводы для такого ограничения совокупности рассматриваемых значений величины x могут быть самые разнообразные, как чисто логические, так и более реальные; отчасти мы уже говорили об этом в лекции I. Если мы, например, определяем величину $f(x)$ как периметр правильного x-угольника, вписанного в окружность радиуса 1, то само собой ясно, что этим условием она определена для всех целых

$x \geqslant 3$ и *только* для них; мы говорим, что $y = f(x)$ есть в этом случае функция от x, определенная на множестве всех целых $x \geqslant 3$. Формально ничто не мешает нам, конечно, доопределить функцию $f(x)$, произвольным образом задавая ее значения для всех остальных значений величины x; но если в этом нет надобности, то мы этого и не делаем, оставляя функцию $f(x)$ никак не определенной вне вышеупомянутого множества. Другой пример: если в данном физическом исследовании x означает температуру какого-либо тела, выраженную в градусах Цельсия, то, вероятно, будет совершенно бессмысленно и бесцельно определять функцию $f(x)$ для значений величины x, меньших чем -273.

Было бы, таким образом, нецелесообразно требовать от каждой функции $y = f(x)$, чтобы она была определена для *всех* значений величины x, ибо, настаивая на этом, мы должны были бы во всех реальных задачах произвольным, искусственным и совершенно ненужным образом расширять естественную область тех значений независимой переменной, для которых представляет интерес рассматривать данную функцию. Это заставляет нас внести необходимое уточнение в самое определение функциональной зависимости.

Условимся говорить, что *величина $y = f(x)$ есть функция величины x, определенная на множестве M вещественных чисел, если каждому значению $x \in M$ соответствует некоторое определенное значение величины y.* Это определение функциональной зависимости, несколько более точное, чем данное нами в начале лекции I, сразу разрешает все сомнения и позволяет нам при исследовании каждой отдельной функции ограничиться тем множеством значений величины x, той естественной «областью определения» функции y, какая диктуется целью данного исследования.

Выбор этой области определения функции может, как мы уже знаем, диктоваться как чисто математическими, так и реальными, например физическими, соображениями. Здесь необходимо только предостеречь вас от одного печального смешения, к сожалению, еще очень часто встречающегося и основанного все на той

же рудиментарной традиции — не делать различия между функцией и аналитическим выражением. Область определения функции часто хотят поставить в зависимость от того, в какой области имеет смысл то или другое аналитическое выражение. Так, говорят, что «областью определения функции $\sqrt{1-x^2}$ является отрезок $[-1, +1]$» или «областью определения функции $\lg x$ является полупрямая $x > 0$». На самом деле здесь речь идет, конечно, не об области определения какой-либо функции, а о той области, в которой имеет смысл данное аналитическое выражение. Так, очень может случиться (математика знает огромное множество примеров этого рода), что нам придется иметь дело с некоторой функцией $y = f(x)$, определенной на отрезке $[0, 2]$ и действительно интересующей нас на всем этом отрезке, и что для $0 \leqslant x \leqslant 1$

$$y = \sqrt{1-x^2};$$

из этого мы никак не можем заключить, что отрезок $[1, 2]$ лежит вне области определения нашей функции и потому нас интересовать не должен; напротив, мы будем искать для отрезка $[1, 2]$ другое аналитическое выражение этой функции (основываясь, конечно, на ее определении); а если нам не удастся такого найти, то будем исследовать ее на этом участке другими, не аналитическими методами.

Область определения каждой функции выбирается на основе соображений либо чисто математических, либо прикладных; но во всех случаях эти соображения исходят из существа дела и не должны идти на поводу у чисто формальных особенностей того или другого специального аналитического аппарата.

Непрерывность. Приступая к изучению функциональных зависимостей, мы должны, конечно, прежде всего с помощью целесообразной классификации внести хотя бы некоторый порядок в предстоящий нам многообразный мир. Первым таким классифицирующим и организующим принципом служит обычно (и с полным основанием) разделение всех функций на *непрерывные* и *разрывные*, причем математический анализ фактически

имеет дело почти исключительно с непрерывными функциями, лишь в сравнительно редких случаях привлекая к рассмотрению и простейшие из разрывных. Непрерывные функции обладают целым рядом особых свойств, которых лишены, вообще говоря, функции разрывные; благодаря этим свойствам исследование и применение непрерывных функций весьма значительно облегчаются, так что изучение этих свойств становится для анализа чрезвычайно важным делом.

Мы говорим, что функция $y = f(x)$ *непрерывна при* $x = a$ (или, короче, *в точке* a), если $\lim\limits_{x \to a} f(x) = f(a)$ или, что в силу определения понятия предела равносильно тому же, если для любой окрестности V числа $f(a)$ найдется такая окрестность U числа a, что для любого $x \in U$ мы имеем $f(x) \in V$. Таким образом, для непрерывности функции в точке a требуется, во-первых, существование предела $\lim\limits_{x \to a} f(x)$ и, во-вторых, совпадение этого предела с тем значением, которое функция принимает при $x = a$. Само собой разумеется, что второе из первого еще не вытекает, как показывает пример функции

$$f(x) = \begin{cases} x^2 & \text{при } x \neq 0, \\ 1 & \text{при } x = 0. \end{cases} \tag{3}$$

По поводу этого определения надо прежде всего заметить, что так понимаемая непрерывность есть *локальное* (местное) свойство функции, т. е. такое свойство, которым функция может обладать в одной точке и не обладать в другой; так, функция (3) разрывна (т. е. не непрерывна) при $x = 0$ и непрерывна при любом другом значении x; это — очень важное обстоятельство, которое никогда не надо упускать из вида.

Далее, мы называем функцию *непрерывной на данном отрезке* $[a, b]$, если она в вышеприведенном смысле непрерывна в каждой точке этого отрезка; при этом в точке a требуется лишь непрерывность *справа*, т. е. соотношение $\lim\limits_{x \to a+0} f(x) = f(a)$, а в точке b — непрерывность *слева*, определяемая аналогичным соотношением, которое вы напишете сами (если имеется в виду *от-*

крытый отрезок (a, b), то, разумеется, в точках a и b от функции ничего не требуется). Заметим кстати, что математики давно уже пользуются очень удобным обозначением

$$\lim_{x \to a+0} f(x) = f(a+0), \quad \lim_{x \to a-0} f(x) = f(a-0),$$

с помощью которого определение непрерывности функции $f(x)$ в точке a можно записать посредством весьма простого соотношения

$$f(a+0) = f(a-0) = f(a);$$

это обозначение не может привести ни к каким смешениям, если только помнить, что $f(a+0)$ и $f(a-0)$ представляют собой не значения функции $f(x)$ в каких-либо точках, а пределы таких значений при некоторых определенных изменениях величины x.

Ограниченные функции. Теперь мы должны ознакомиться с другим свойством функций, которое в противоположность свойству непрерывности является не локальным, а *интегральным*, т. е. может быть определено сразу для любого множества значений независимой переменной, не нуждаясь в предварительном определении для отдельных ее значений (в отдельных точках).

Функция $y = f(x)$ называется *ограниченной на множестве* M, если все значения, принимаемые ею на этом множестве, принадлежат некоторому закрытому отрезку; очевидно, вместо этого мы можем предъявить и совершенно равносильное требование: существует такое положительное число c, что $|f(x)| < c$ для всех $x \in M$. Более детально, мы называем функцию y *ограниченной сверху* (*снизу*) на M, если существует такое число c, что $f(x) < c$ $(f(x) > c)$ для всех $x \in M$, Чтобы функция была просто ограниченной, она должна быть, очевидно, ограничена как сверху, так и снизу.

Свойство ограниченности не есть, подобно свойству непрерывности на отрезке, совместное выполнение известных требований в каждой точке этого отрезка. Число c, о котором говорится в определении ограниченности, выбирается сразу для всего множества M.

В каждой отдельной точке этого множества, если функция в ней определена, такое число c существует тривиальным образом: для точки x достаточно положить, например, $c = |f(x)| + 1$. Но функция, определенная, например, в каждой точке некоторого отрезка, может быть и неограниченной на этом отрезке; чтобы в этом убедиться, вспомним, что $\operatorname{tg} x$ возрастает безгранично при $x \to \frac{\pi}{2} - 0$, так что функция

$$y = \begin{cases} \operatorname{tg} x \left(0 \leqslant x < \frac{\pi}{2}\right), \\ 0 \qquad \left(x = \frac{\pi}{2}\right) \end{cases}$$

не ограничена на отрезке $\left[0, \frac{\pi}{2}\right]$.

Как для многих интегральных свойств, можно, однако, и для ограниченности функции на данном отрезке указать такое локальное свойство, выполнение которого в каждой точке данного отрезка равносильно выполнению рассматриваемого интегрального свойства. Условимся называть функцию y *ограниченной в точке* x, если она ограничена в некоторой окрестности U числа x (обратите внимание, что здесь локальное свойство определяется через ранее определенное интегральное; в случае непрерывности дело обстояло как раз наоборот). Мы можем теперь утверждать, что *для ограниченности функции* $y = f(x)$ *на отрезке* $[a, b]$ (*закрытом*) *необходимо и достаточно, чтобы она была ограничена в каждой точке этого отрезка.* Необходимость этого условия вытекает из самих определений и не нуждается в доказательстве; чтобы показать его достаточность, допустим, что каждое число x отрезка $[a, b]$ может быть окружено окрестностью U_x, в которой функция y ограничена; применяя лемму Гейне — Бореля, мы находим, что отрезок $[a, b]$ покрывается конечным числом отрезков $\Delta_1, \Delta_2, \ldots, \Delta_n$, на каждом из которых y ограничена. Если $|y| < c_i$ на отрезке Δ_i $(i = 1, 2, \ldots, n)$ и если c есть наибольшее из чисел $c_1, c_2, \ldots, c_n$, то $|y| < c$ для всех $x \in [a, b]$, чем наше утверждение и доказано.

Условимся называть множество чисел N *ограниченным*, если все входящие в него числа могут быть заключены в некоторый закрытый отрезок. Очевидно, что ограниченность функции $y = f(x)$ на множестве M равносильна ограниченности множества N значений, принимаемых этой функцией, когда величина x «пробегает» множество M, т. е. принимает всевозможные значения, принадлежащие этому множеству. Само собою ясно, что означают термины «множество N ограничено сверху (или справа)» и «множество N ограничено снизу (или слева)».

Условимся называть число β *верхней гранью* множества N, если: 1) множество N не содержит чисел, бо́льших чем β, и 2) в любой окрестности числа β найдется число, принадлежащее этому множеству. Подобным же образом *нижней гранью* множества N мы назовем такое число α, что: 1) в множестве N нет чисел, меньших чем α, и 2) в любой окрестности числа α найдется число, принадлежащее множеству N. Очевидно, что множество, имеющее верхнюю (нижнюю) грань, ограничено сверху (снизу). В анализе играет значительную роль обратное предложение:

Теорема. *Всякое непустое множество чисел, ограниченное сверху (снизу), имеет единственную верхнюю (нижнюю) грань.*

В частности, всякое ограниченное множество имеет как верхнюю, так и нижнюю грань.

Доказательство. Разобьем все вещественные числа на два класса A и B по следующему принципу: $x \in A$, если вправо от x имеется хотя бы одна точка множества N, и $x \in B$ — в противном случае. Вы сами без труда убедитесь, что это разбиение есть сечение; пусть α — рубеж этого сечения; покажем, что α есть верхняя грань множества N.

Прежде всего убедимся, что вправо от α не может быть точек множества N; в самом деле, если $\beta \in N$ и $\beta > \alpha$, то, полагая $\gamma = \frac{1}{2}(\alpha + \beta)$, мы имеем $\alpha < \gamma < \beta$; из $\gamma > \alpha$ следует $\gamma \in B$, а из $\beta > \gamma$ и $\beta \in N$ следует $\gamma \in A$, и, таким образом, мы пришли к противоречию.

Далее, пусть $U = (\alpha_1, \alpha_2)$ $(\alpha_1 < \alpha < \alpha_2)$ — любая окрестность точки α.

Очевидно, что $\alpha_1 \in A$, $\alpha_2 \in B$; в силу первого из этих соотношений имеются точки множества N вправо от α_1, в силу же второго соотношения таких точек нет вправо от α_2; значит, такие точки имеются в отрезке (α_1, α_2). Число α обладает, таким образом, обоими свойствами верхней грани, существование которой тем самым доказано.

Эта грань единственна. Действительно, если бы множество N имело две верхние грани β и β' $(\beta < \beta')$, то мы немедленно получили бы противоречие, замечая, что, с одной стороны, множество N не может содержать чисел $> \beta$ (так как β есть его верхняя грань), а с другой стороны, оно должно содержать такие числа в силу того, что их должна содержать сколь угодно малая окрестность числа β'. Очевидно, что для случая нижней грани теорема доказывается совершенно аналогично.

Наглядно мы должны представлять себе грани данного ограниченного множества как концы наименьшего отрезка, содержащего все числа этого множества. Очевидно, что мы можем определить верхнюю грань данного множества и как наименьшее из тех чисел c, по отношению к которому все числа $x \in N$ удовлетворяют неравенству $x \leqslant c$, и что аналогичное определение возможно и для нижней грани.

Очень важно иметь в виду, что каждая из граней данного ограниченного множества N может как принадлежать, так и не принадлежать этому множеству. Так, закрытый отрезок содержит свои грани, а открытый их не содержит; множество чисел вида $\frac{1}{n}$, где n — любое натуральное число, содержит свою верхнюю грань (число 1), но не содержит своей нижней грани (числа 0).

Если функция $y = f(x)$ ограничена на множестве M, то, как мы уже заметили, множество N значений, принимаемых ею на этом множестве, ограничено и, следовательно, имеет, в силу доказанной теоремы, верхнюю и нижнюю грани. Эти грани множества N мы условимся называть *гранями функции y на множестве*

M. Таким образом, *всякая функция, ограниченная на данном множестве, имеет на нем единственную верхнюю и единственную нижнюю грань.*

Каждая из этих граней может быть некоторым значением данной функции; тогда она служит наибольшим или наименьшим значением этой функции на данном множестве; но может случиться, что та или другая из граней вовсе не принадлежит совокупности значений функции на данном множестве; в таком случае функция на этом множестве вообще не имеет наибольшего (или наименьшего) значения. Так, функция

$$y = \begin{cases} x \text{ при } 0 \leqslant x < 1, \\ 0 \text{ при } x = 1 \end{cases} \tag{4}$$

имеет, очевидно, грани 0 и 1 на отрезке $[0, 1]$; нижняя грань является (наименьшим) значением этой функции (при $x = 0$ и $x = 1$); верхняя же грань — число 1 — среди ее значений не встречается; функция не имеет наибольшего значения на отрезке $[0, 1]$.

Основные свойства непрерывных функций. Теперь мы установим четыре важнейших свойства непрерывных функций.

Лемма. *Если функция $y = f(x)$ непрерывна в точке a и если $f(a) < b$, то существует такая окрестность U точки a, что для всех $x \in U$ мы имеем $f(x) < b$.*

Это — почти тривиальный вывод из самого определения непрерывности функции в точке. В самом деле, если числа α и β выбраны так, что $\alpha < f(a) < \beta < b$, то в силу этого определения для всех x некоторой окрестности числа a мы должны иметь

$$\alpha < f(x) < \beta < b,$$

что и требовалось доказать.

Само собою ясно, что лемма остается верной, если всюду в ее формулировке знак $<$ мы заменим знаком $>$.

Теорема 1. *Функция $y = f(x)$, непрерывная на закрытом отрезке $[a, b]$, ограничена на этом отрезке.*

Для доказательства заметим, что функция y, будучи непрерывной на отрезке $[a, b]$, непрерывна в каждой

точке этого отрезка; пусть λ — любая точка отрезка $[a, b]$; по самому определению непрерывности существует такая окрестность U числа λ, что для любого $x \in U$

$$f(\lambda) - 1 < f(x) < f(\lambda) + 1;$$

но это показывает, что функция $f(x)$ ограничена в окрестности U (все ее значения содержатся в некотором отрезке); значит, она ограничена *в точке* λ (ибо ограниченность в точке нами определена как ограниченность в некоторой окрестности этой точки). Будучи же ограниченной в любой точке отрезка $[a, b]$, функция y, как мы видели выше, должна быть ограниченной и на всем этом отрезке.

Теорема 2. *Функция $y = f(x)$, непрерывная на закрытом отрезке $[a, b]$, имеет на нем наибольшее и наименьшее значения.*

Доказательство. В силу теоремы 1 функция y ограничена на отрезке $[a, b]$ и, значит, имеет на нем нижнюю грань α и верхнюю грань β; надо только показать, что эти грани являются значениями функции y. Покажем это, например, для верхней грани β.

Если бы β не было значением функции y на отрезке $[a, b]$, то это означало бы, что $f(x) < \beta$ для всех чисел x этого отрезка. Пусть β_x — любое число, заключенное между $f(x)$ и β, так что $f(x) < \beta_x < \beta$. В силу нашей леммы мы можем окружить точку x окрестностью U_x, для любой точки x' которой $f(x') < \beta_x$. Система окрестностей U_x, построенных для всех точек x отрезка $[a, b]$, покрывает этот отрезок и в силу теоремы Гейне — Бореля содержит конечную подсистему окрестностей $U_{x_1}, U_{x_2}, \ldots, U_{x_n}$, также покрывающую отрезок $[a, b]$. Но для любого числа x' отрезка U_{x_k} мы имеем $f(x') < \beta_{x_k}$; обозначая через β' наибольшее из чисел $\beta_{x_1}, \beta_{x_2}, \ldots, \beta_{x_n}$, мы видим, что для всех $x \in [a, b]$

$$f(x) < \beta' < \beta,$$

вследствие чего β не может быть верхней гранью функции $f(x)$ на отрезке $[a, b]$. Полученное противоречие и доказывает теорему.

Выше [см. (4)] мы видели пример ограниченной функции, не имеющей на данном отрезке наибольшего значения; мы знаем теперь, что это возможно лишь для разрывных функций; и действительно, функция (4) разрывна при $x = 1$.

Важно еще подчеркнуть, что теорема 2 справедлива только для закрытых отрезков; так, даже столь простые функции, как $y = x$ и $y = x^3$, лишены наибольших и наименьших значений на открытых отрезках.

Теорема 3. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и если $f(a) < \mu < f(b)$ или $f(a) > \mu > f(b)$, то найдется такое число c на отрезке $[a, b]$, что $f(c) = \mu$.*

Короче говоря, непрерывная функция между любыми двумя своими значениями обязательно проходит через все промежуточные значения. Разумеется, разрывные функции, вообще говоря, не обладают этим свойством; так, функция Дирихле на любом отрезке принимает значения 0 и 1, но нигде не принимает ни одного промежуточного между ними значения; для функции (4) мы имеем $f\left(\frac{1}{2}\right) = \frac{1}{2}$, $f(1) = 0$, ни для одного x отрезка $\left(\frac{1}{2}, 1\right)$ функция $f(x)$ не принимает промежуточного значения $\frac{1}{4}$.

Доказательство теоремы 3. Пусть сначала $\mu = 0$, $f(a) < 0$, $f(b) > 0$; требуется доказать, что $f(x)$ обращается в нуль в некоторой внутренней точке отрезка $[a, b]$. Допустим, что это не так, и разделим отрезок $[a, b]$ пополам. Очевидно, что в одной из двух половин $f(x)$ имеет разные знаки в концах половины; эту половину мы снова делим пополам и снова выбираем ту половину, в концах которой $f(x)$ имеет разные знаки, и т. д. Мы получаем, таким образом, стягивающуюся систему отрезков; пусть α — их общая точка; по предположению $f(\alpha) \neq 0$, но если $f(\alpha) > 0$, то в силу леммы (стр. 75) мы имеем $f(x) > 0$ для всех x некоторой окрестности U точки α; это же невозможно, так как U содержит бесчисленное множество отрезков нашей стягивающейся системы, в концах каждого из

которых $f(x)$ принимает значения разных знаков. Тем же путем мы убеждаемся, что $f(\alpha) < 0$ невозможно. Полученное противоречие и доказывает наше утверждение.

Наконец, если $\mu \neq 0$, то для доказательства теоремы 3 достаточно применить только что полученный нами результат к функции $f(x) - \mu$.

Прежде чем формулировать четвертое свойство непрерывных функций, мы должны ввести одно новое понятие, имеющее для изучения этого класса функций важнейшее значение. Вы помните, конечно, что непрерывность есть локальное свойство функций; это положение нисколько не меняется от того, что мы говорим о «функции, непрерывной на данном отрезке», ибо непрерывность на отрезке означает не больше, чем непрерывность в каждой точке этого отрезка, и потому ни в какой мере не изменяет локального характера этого понятия. Можно, однако, выразить идею непрерывности функции на данном отрезке и некоторым другим образом, причем она получает *интегральный* характер, т. е. непосредственно не исходит ни из каких локальных свойств, а характеризует функцию сразу во всем отрезке.

Условимся называть функцию $y = f(x)$ *равномерно непрерывной* на отрезке $[a, b]$, если она обладает следующим свойством: *сколь бы мало ни было положительное число* ε, *существует такое другое положительное число* δ, *что для любых двух чисел* x_1 *и* x_2 *отрезка* $[a, b]$, *отстоящих друг от друга менее чем на* δ,

$$|f(x_1) - f(x_2)| < \varepsilon.$$

Вы видите, что это определение не приурочено ни к какой отдельной точке, а стремится охарактеризовать поведение функции на всем отрезке $[a, b]$, говоря, что в двух точках, достаточно близких друг к другу, значения функции должны быть также сколь угодно близки друг к другу. Понятно, почему этот вид непрерывности называют «равномерным»: вся специфика его именно в том, что здесь требуется некая одинаковость в поведении функции во всех местах данного отрезка; точки x_1 и x_2 могут быть взяты где угодно на отрезке

$[a, b]$, лишь бы их взаимное расстояние не превышало числа δ.

Совершенно ясно, что функция, равномерно непрерывная на отрезке $[a, b]$, по необходимости непрерывна в каждой точке этого отрезка (а значит, и на всем отрезке), ибо в силу равномерной непрерывности из $|x-\alpha|<\delta$ вытекает $|f(x)-f(\alpha)|<\varepsilon$, т. е. для любой окрестности $V=(f(\alpha)-\varepsilon, f(\alpha)+\varepsilon)$ числа $f(\alpha)$ найдется такая окрестность $U=(\alpha-\delta, \alpha+\delta)$ числа α, что из $x\in U$ вытекает $f(x)\in V$; а это и значит, что функция $f(x)$ непрерывна при $x=\alpha$. Чрезвычайно важно, что имеет место и обратное предложение, т. е. что из непрерывности функции y в любой точке (закрытого) отрезка $[a, b]$ вытекает ее равномерная непрерывность на этом отрезке; требование равномерной непрерывности не сужает, таким образом, класса непрерывных функций (поскольку речь идет о закрытом отрезке).

Теорема 4. *Функция $y=f(x)$, непрерывная в каждой точке закрытого отрезка $[a, b]$, равномерно непрерывна на этом отрезке.*

Для открытых отрезков теорема неверна. Так, функция $y=\sin\frac{1}{x}$, очевидно, непрерывная в каждой точке *открытого* отрезка $(0, 1)$, не может быть равномерно непрерывной на этом отрезке, так как, сколь бы мало ни было $\delta>0$, можно найти (вблизи нуля) две точки, отстоящие друг от друга менее чем на δ и такие, что разность значений функции y в них больше, чем 1.

Доказательство теоремы 4. Пусть функция y непрерывна в каждой точке отрезка $[a, b]$; тогда для любой точки x этого отрезка найдется такое число $\delta_x>0$, что $|f(x_1)-f(x_2)|<\varepsilon$, если x_1 и x_2 принадлежат отрезку $[x-\delta_x, x+\delta_x]$. Обозначим через Δ_x отрезок $\left[x-\frac{1}{2}\delta_x, x+\frac{1}{2}\delta_x\right]$. Очевидно, совокупность всех отрезков Δ_x покрывает отрезок $[a, b]$. Поэтому на основании леммы Гейне — Бореля существует конечная система M отрезков Δ_x, также покрывающая отрезок $[a, b]$. Пусть δ — длина наименьшего из отрезков системы M, и пусть x_1 и x_2 — две любые точки отрезка

$[a, b]$, отстоящие друг от друга меньше чем на $\frac{1}{2}\delta$. Я утверждаю, что $|f(x_1)-f(x_2)|<\varepsilon$ (чем теорема 4 и будет доказана). В самом деле, точка x_1 принадлежит какому-нибудь отрезку $\Delta_x=\left[x-\frac{1}{2}\delta_x,\ x+\frac{1}{2}\delta_x\right]$ системы M; но тогда точка x_2, отстоящая от x_1 менее чем на $\frac{1}{2}\delta\leqslant\frac{1}{2}\delta_x$, принадлежит отрезку $[x-\delta_x, x+\delta_x]$, в который входит, конечно, и x_1; отсюда, по самому определению числа δ_x, мы действительно имеем $|f(x_1)-f(x_2)|<\varepsilon$.

Непрерывность элементарных функций. Вы, конечно, знаете, что алгебраическая сумма и произведение любого числа непрерывных функций также являются непрерывными функциями; отношение двух непрерывных функций непрерывно в каждом отрезке, в котором знаменатель не обращается в нуль; все эти предложения справедливы как в локальном, так и в интегральном понимании, т. е. независимо от того, имеем ли мы в виду непрерывность в точке или на отрезке или, наконец, равномерную непрерывность. Доказательства этих предложений можно найти в любом курсе анализа; они не представляют принципиального интереса, и здесь нам нет надобности на них останавливаться.

Значительно интереснее вопрос о непрерывности так называемых элементарных функций, т. е. того небольшого класса функций, с которыми оперирует элементарная математика и которые и в высшей математике полностью сохраняют свое ведущее значение. Сюда относятся прежде всего все функции, значения которых получаются из значений независимой переменной применением шести основных алгебраических операций, а затем — небольшое число трансцендентных функций: тригонометрические функции (прямые и обратные), показательные и логарифмические функции; сюда же относятся, наконец, и всевозможные конечные комбинации перечисленных видов функций, например $x^2(1-x\sin x)$ или $\cos 2^{x^2}$.

Все элементарные функции в основном непрерывны; лишь некоторые из них претерпевают разрывы в

некоторых отдельных точках $\left[\frac{1}{x}\right.$ при $x=0$, $\operatorname{tg} x$ при $x=\frac{\pi}{2}$ и т. д.[1]$\left.\right]$. Не для всех элементарных функций, однако, доказательство этого факта легко; в курсах анализа обычно этот вопрос не рассматривается во всей полноте; между тем он имеет важное значение, и нам придется сделать по этому поводу несколько замечаний.

Прежде всего, очень легко устанавливается непрерывность всех многочленов при любых значениях независимой переменной; в самом деле, так как для любого многочлена $P(x)$ и для любого числа a разность $P(x)-P(a)$ алгебраически делится без остатка на разность $x-a$ (теорема Безу), то

$$P(x)-P(a)=(x-a)Q(x),$$

где $Q(x)$ — многочлен; отсюда при $x \rightarrow a$, так как многочлен $Q(x)$, очевидно, ограничен в точке a, мы имеем $P(x) \rightarrow P(a)$, т. е. многочлен $P(x)$ непрерывен в точке a. Далее, из непрерывности многочленов непосредственно вытекает непрерывность всякой рациональной дроби $\frac{P_1(x)}{P_2(x)}$ в любом отрезке, на котором $P_2(x)$ ни в одной точке не обращается в нуль.

Столь же просто вопрос решается и для тригонометрических функций. Формула

$$\sin x-\sin a=2 \cos \frac{x+a}{2} \sin \frac{x-a}{2}$$

и вытекающее из нее неравенство

$$|\sin x-\sin a| \leqslant 2\left|\sin \frac{x-a}{2}\right| \leqslant 2\left|\frac{x-a}{2}\right|=|x-a|$$

показывают, что $\sin x \rightarrow \sin a$ при $x \rightarrow a$, т. е. что синус — непрерывная функция. Совершенно аналогично

---

[1]) Заметим, что функции $\frac{1}{x}$ и $\operatorname{tg} x$ непрерывны во всей области своего определения. В точках разрыва, о которых идет речь, функции эти не определены. О таких точках разрыва см. стр. 85.

устанавливается непрерывность косинуса. Отсюда функция

$$y = \operatorname{tg} x = \frac{\sin x}{\cos x}$$

непрерывна при всех значениях x, для которых $\cos x \neq 0$, т. е. для значений вида $x \neq (2k+1)\frac{\pi}{2}$, где k — любое целое число.

Сложнее доказательство непрерывности показательной функции a^x. Допустим для определенности, что $a > 1$, так что функция $y = a^x$ возрастающая. Покажем сперва, что $a^x \to a^0 = 1$, когда $x \to 0$. Снова для определенности допустим, что $x \to +0$, т. е. будем рассматривать a^x лишь для положительных значений x; случай $x \to -0$ разрешается совершенно аналогично. При $x > 0$ всегда $a^x > 1$, и, чтобы доказать, что при $x \to +0$ функция $a^x \to 1$, надо только убедиться, что $a^x < 1 + \varepsilon$ при достаточно малых x, сколь бы мало ни было $\varepsilon > 0$. Но для любого натурального числа n

$$(1+\varepsilon)^n = 1 + n\varepsilon + \ldots > n\varepsilon,$$

откуда $(1+\varepsilon)^n \to +\infty$ при $n \to +\infty$. Значит, при достаточно большом n

$$(1+\varepsilon)^n > a, \quad a^{1/n} < 1 + \varepsilon;$$

а так как a^x — возрастающая функция, то и для всех достаточно малых x

$$1 < a^x < 1 + \varepsilon,$$

т. е. действительно $a^x \to 1$ при $x \to 0$. Дальнейшее уже просто. Соотношение

$$a^x - a^\alpha = a^\alpha (a^{x-\alpha} - 1)$$

показывает, что при $x \to \alpha$ функция a^x имеет пределом a^α, т. е. что эта функция непрерывна при всех значениях x.

Вопрос о непрерывности иррациональных, логарифмических и обратных тригонометрических функций может теперь быть решен на основании общей теоремы о непрерывности обратных функций: *если функция $y = f(x)$ — непрерывная и возрастающая на отрезке*

$[a, b]$, *то обратная функция* $x = \varphi(y)$ *также непрерывна на отрезке* $[\alpha, \beta]$, *где* $\alpha = f(a)$, $\beta = f(b)$. В самом деле, пусть γ — любая точка отрезка $[\alpha, \beta]$. Надо доказать, что

$$\lim_{y \to \gamma} \varphi(y) = \varphi(\gamma).$$

Пусть сначала $y \to \gamma - 0$, т. е. y приближается к γ снизу; так как вместе с функцией $f(x)$ и обратная функция $\varphi(y)$ также возрастающая, то $\lim\limits_{y \to \gamma - 0} \varphi(y)$ во всяком случае существует; обозначим этот предел через c и покажем, что $c = \varphi(\gamma)$.

Так как функция $f(x)$ непрерывна и $\varphi(y) \to c$ при $y \to \gamma - 0$, то $y = f[\varphi(y)] \to f(c)$ при $y \to \gamma - 0$, откуда $f(c) = \gamma$; но это непосредственно дает: $c = \varphi[f(c)] = \varphi(\gamma)$, т. е.

$$\lim_{y \to \gamma - 0} \varphi(y) = \varphi(\gamma);$$

а так как, очевидно, мы тем же путем легко убеждаемся, что и

$$\lim_{y \to \gamma + 0} \varphi(y) = \varphi(\gamma),$$

то

$$\lim_{y \to \gamma} \varphi(y) = \varphi(\gamma),$$

что и требовалось доказать.

Для того чтобы иметь возможность без специального исследования утверждать непрерывность таких комбинированных элементарных функций, как $\lg \sin x$, $2^{3x^2 - 4x}$ и т. п., нужна еще следующая «теорема о непрерывности функции от функции»:

Если $y = f(x)$ — *функция, непрерывная при* $a \leqslant x \leqslant b$, *а* $z = \varphi(y)$ — *функция, непрерывная при* $\alpha \leqslant y \leqslant \beta$, *где* α *и* β *суть соответственно наименьшее и наибольшее значения функции* $f(x)$ *на отрезке* $[a, b]$, *то функция* $z = \varphi[f(x)]$ *непрерывна на отрезке* $[a, b]$.

Доказательство. Пусть c — любая точка отрезка $[a, b]$, $f(c) = \gamma$, $\varphi(\gamma) = \varphi[f(c)] = \zeta$; пусть W — любая окрестность точки ζ; в силу непрерывности функции $\varphi(y)$ существует такая окрестность V точки γ, что $\varphi(y) \in W$ при $y \in V$; наконец, в силу не-

прерывности функции $f(x)$ существует такая окрестность U точки c, что при $x \in U$ мы имеем $f(x) \in V$, а значит, $\varphi[f(x)] \in W$. Так как W есть любая окрестность точки $\zeta = f[\varphi(c)]$, то отсюда следует

$$\lim_{x \to c} f[\varphi(x)] = f[\varphi(c)],$$

т. е. непрерывность функции $f[\varphi(x)]$ в произвольной точке c отрезка $[a, b]$ доказана.

Колебание функции в точке. Мы получаем наиболее удобный подход к изучению разрывных функций, вводя понятия колебания функции на данном отрезке и в данной точке. Пусть на отрезке $[a, b]$, за исключением, быть может, конечного числа его точек, определена совершенно произвольная функция $f(x)$. Если она на данном отрезке не ограничена, мы будем говорить, что колебание ее на этом отрезке равно $+\infty$; если же ограничена, то колебанием ее на отрезке $[a, b]$ назовем разность $M-m$ между ее верхней гранью M и нижней гранью m на этом отрезке; во всех случаях будем обозначать колебание функции $f(x)$ на отрезке $[a, b]$ символом $\omega_f(a, b)$.

Если отрезок $[a', b']$ составляет часть отрезка $[a, b]$ (т. е. если $a \leqslant a' < b' \leqslant b$), то, очевидно, грани M', m' функции $f(x)$ на отрезке $[a', b']$ подчиняются неравенствам $m \leqslant m' \leqslant M' \leqslant M$ и, следовательно, $\omega_f(a', b') \leqslant \omega_f(a, b)$. Поэтому, если мы выберем на отрезке $[a, b]$ произвольную точку c [1]), окружим ее замкнутой окрестностью $[\alpha, \beta]$ и заставим концы α и β этой окрестности приближаться к точке c, то $\omega_f(\alpha, \beta)$, изменяясь при этом приближении, никогда не может возрастать, а либо убывает, либо остается постоянной; а так как колебание $\omega_f(\alpha, \beta)$, будучи неотрицательной величиной, ограничено снизу, то в силу леммы 1′ лекции I (стр. 31) оно должно при этом стремиться к некоторому пределу, который мы будем обозначать через $\omega_f(c)$ и называть *колебанием функции $f(x)$ в точке c*. Таким образом,

$$\omega_f(c) = \lim_{\substack{\alpha \to c-0 \\ \beta \to c+0}} \omega_f(\alpha, \beta).$$

[1]) В этой точке функция может быть и не определена.

Введение этого понятия позволяет прежде всего поновому взглянуть на непрерывность функции в данной точке, как показывает следующая

Теорема. *Для того чтобы функция* $f(x)$, *определенная в некоторой окрестности точки* c, *была непрерывна в этой точке* c, *необходимо и достаточно, чтобы* $\omega_f(c)=0$.

Доказательство. 1) Пусть $\omega_f(c)=0$. Это значит, что в достаточно малой окрестности $U=(\alpha, \beta)$ точки c колебание функции $f(c)$ становится как угодно малым. А так как, очевидно, для любого $x\in U$

$$|f(x)-f(c)|\leqslant M-m=\omega_f(\alpha, \beta)$$

(где M и m — грани функции $f(x)$ на отрезке $[\alpha, \beta]$), то и подавно $|f(x)-f(c)|$ как угодно мало при $x\in U$; но это означает, что $f(x)$ непрерывна в точке c.

2) Если $\omega(c)=\omega>0$, то, какова бы ни была окрестность $U=(a, b)$ точки c, $M-m\geqslant\omega$; поэтому существуют такие числа $\alpha\in[a, b]$ и $\beta\in[a, b]$, что $f(\beta)-f(\alpha)>\frac{\omega}{2}$, но

$$f(\beta)-f(\alpha)=[f(\beta)-f(c)]+[f(c)-f(\alpha)];$$

из того, что $f(\beta)-f(\alpha)\geqslant\frac{\omega}{2}$, вытекает, что по меньшей мере одно из двух слагаемых правой части должно быть $\geqslant\frac{\omega}{4}$; таким образом, существует такая точка x отрезка $[a, b]$, что $|f(x)-f(c)|\geqslant\frac{\omega}{4}$; так как окрестность U произвольна, то функция $f(x)$ не может быть непрерывной в точке c.

Если функция имеет в некоторой точке (в которой она может быть и не определена) колебание, большее нуля, то говорят, что она разрывна в этой точке. Точки, в которых данная функция непрерывна, называют для краткости ее *точками непрерывности*, а точки, в которых она разрывна,— *точками разрыва*. Мы видим, что $\omega_f(c)=0$ в точках непрерывности и $\omega_f(c)>0$ в точках разрыва.

Точки разрыва. Уже это обстоятельство дает нам повод сделать попытку внесения некоторого порядка в

чрезвычайно разнообразную совокупность возможных типов точек разрыва. Раз условие $\omega(c) > 0$ отличает точки разрыва от точек непрерывности, то мы, естественно, ожидаем, что величина $\omega_f(c)$, т. е. колебание функции в данной точке, сможет служить разумной и удобной *мерой* разрывности функции в данной точке. Мы еще более укрепимся в этом ожидании, если вспомним определение величины $\omega_f(c)$; ведь $\omega_f(\alpha, \beta)$ является мерой взаимного расхождения значений функции $f(x)$ на отрезке $[\alpha, \beta]$; предел этой величины при условии, что отрезок $[\alpha, \beta]$ стягивается к точке c, может поэтому рассматриваться нами как мера взаимного расхождения значений функции в точках, сколь угодно близких к c, т. е. мера именно той особенности этой функции, которая обусловливает претерпеваемый ею в точке c разрыв. Условимся, таким образом, называть величину $\omega_f(c)$ *мерой разрыва* функции $f(x)$ в точке c; это дает нам возможность сравнивать между собой различные точки разрыва по той «степени разрывности», какую в них обнаруживает данная функция; в частности, наибольшую разрывность функция имеет, конечно, в тех точках, в которых она не ограничена ($\omega_f(c) = +\infty$).

С понятием меры разрыва (или колебания функции в точке) связано одно важное предложение, служащее непосредственным обобщением теоремы 4 о равномерной непрерывности (см. стр. 79).

Теорема. *Если мера разрыва (т. е. колебание) функции $f(x)$ в каждой точке отрезка $[a, b]$ не превосходит числа $\lambda \geqslant 0$, то, сколь бы мало ни было положительное число ε, найдется такое другое положительное число δ, что колебание функции $f(x)$ на любом отрезке длины $< \delta$ не превосходит $\lambda + \varepsilon$.*

Такую функцию можно было бы назвать «непрерывной с точностью до λ»; при $\lambda = 0$ мы получаем, в частности, как раз теорему 4. Вы без всяких затруднений сможете самостоятельно провести доказательство новой теоремы по той же схеме, какой мы пользовались в случае теоремы 4; все отличие будет состоять в том, что роль прежнего ε теперь будет играть число $\lambda + \varepsilon$.

Вы видите, таким образом, что не только локальная непрерывность, но и локальное свойство, состоящее в ограничении меры разрыва функции в данной точке, имеет свой интегральный эквивалент. Доказанная нами теорема имеет целый ряд приложений, в частности, в интегральном исчислении; мы встретимся еще с нею в одной из следующих лекций.

Весьма важным для точек разрыва оказывается и другой классификационный принцип, в основном позволяющий сравнивать их между собой не по величине, а, так сказать, по форме разрыва. Как мы знаем, критерием непрерывности функции $f(x)$ в точке a могут служить равенства $f(a+0)=f(a)=f(a-0)$; может случиться, что $f(a+0)$ и $f(a-0)$ существуют, но по меньшей мере одно из этих двух чисел отлично от $f(a)$; в этом случае мы называем точку a точкой разрыва *первого рода* для функции $f(x)$. Таким образом, точка разрыва первого рода характеризуется тем, что при подходе к ней как справа, так и слева существует предел функции $f(x)$, но что либо эти два предела различны между собою, либо, совпадая между собою, они отличны от значения функции в этой точке. Пример первого типа дает (схематически) рис. 13 (здесь значение функции в точке разрыва может быть каким угодно), пример второго типа — рис. 14.

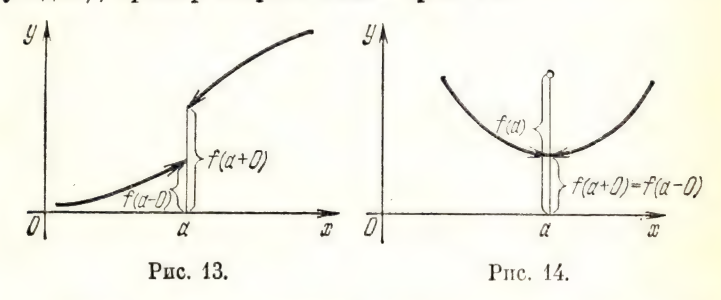


Рис. 13. Рис. 14.

Точки разрыва первого рода представляют собой простейший тип точек разрыва; сравнительная легкость их изучения обусловливается, разумеется, существованием пределов $f(a+0)$ и $f(a-0)$. Все другие точки разрыва называют точками разрыва *второго*

рода. Значит, в точке разрыва второго рода функция не стремится ни к какому пределу по меньшей мере при одном из двух подходов (справа или слева) к этой точке; пример такого поведения дает неоднократно уже упоминавшаяся нами, весьма поучительная функция $y = \sin \frac{1}{x}$ при $x = 0$ (рис. 15). Здесь на любом отрезке, содержащем 0, верхняя грань функции равна $+1$,

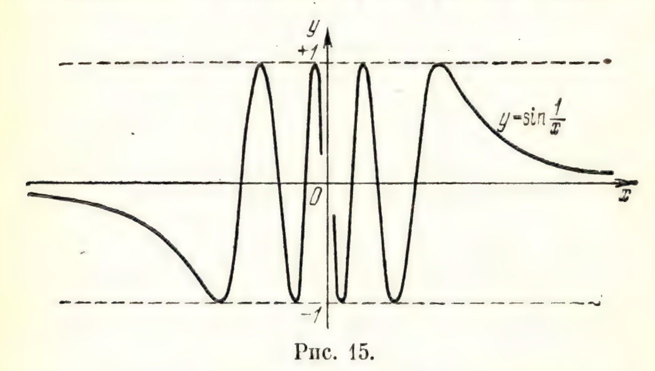


Рис. 15.

а нижняя грань функции равна -1, поэтому $\omega_f(0) = 2$. Функция Дирихле, которую мы рассматривали в начале этой лекции, в каждой точке имеет, очевидно, разрыв второго рода.

Монотонные функции. Функция $y = f(x)$ называется *неубывающей* на отрезке $[a, b]$, если при $a \leqslant x_1 \leqslant x_2 \leqslant b$ всегда $f(x_1) \leqslant f(x_2)$; если при том же условии всегда $f(x_1) \geqslant f(x_2)$, функция $f(x)$ называется *невозрастающей* на отрезке $[a, b]$. Неубывающие и невозрастающие функции вместе образуют класс *монотонных* функций. Монотонные функции обладают целым рядом специальных свойств, которые делают их во многих случаях очень удобным орудием исследования.

Прежде всего, всякая функция $f(x)$, монотонная на данном отрезке $[a, b]$, ограничена на этом отрезке [как обычно, отрезок предполагается закрытым; для открытых отрезков утверждение неверно: функция $y = \frac{1}{x}$ монотонна, но не ограничена на открытом от-

резке $(0, 1)]$; в самом деле, при $a \leqslant x \leqslant b$ значение $f(x)$ заключено между $f(a)$ и $f(b)$; очевидно, далее, что гранями монотонной функции служат ее значения $f(a)$ и $f(b)$ в концах данного отрезка; эти же числа служат наибольшим и наименьшим значениями монотонной функции $f(x)$ на отрезке $[a, b]$.

Монотонная функция может иметь точки разрыва, как нас учит один взгляд на рис. 13. Однако разрывы такой функции ограничены как по своему характеру, так и по своему числу. Прежде всего, в силу ограниченности монотонные функции не могут иметь разрывов бесконечной меры. Далее, *число точек, в которых мера разрыва монотонной функции $f(x)$ превосходит какое-либо положительное число τ, на отрезке $[a, b]$ не больше, чем $\frac{f(b)-f(a)}{\tau}$*. В самом деле, пусть для определенности $f(b) > f(a)$ и пусть функция $f(x)$ имеет n точек на отрезке $[a, b]$, в каждой из которых мера разрыва этой функции превосходит τ; обозначим эти n точек, считая слева направо, через $c_1, c_2, \ldots, c_n$, так что при сколь угодно малом $\varepsilon > 0$

$$f(c_k+\varepsilon)-f(c_k-\varepsilon) > \tau \quad (1 \leqslant k \leqslant n);$$

если мы (что всегда возможно) выберем ε столь малым, чтобы $a < c_1 - \varepsilon$, $c_n + \varepsilon < b$ и

$$c_k+\varepsilon < c_{k+1}-\varepsilon \quad (1 \leqslant k \leqslant n-1),$$

т. е. чтобы окрестности $(c_k-\varepsilon,\ c_k+\varepsilon)$ попарно не имели общих точек и целиком умещались в отрезке $[a, b]$, то, очевидно,

$$n\tau < \sum_{k=1}^{n} [f(c_k+\varepsilon)-f(c_k-\varepsilon)] \leqslant f(b)-f(a),$$

откуда

$$n < \frac{f(b)-f(a)}{\tau},$$

что и требовалось доказать. В частности, таким образом, монотонная функция может иметь лишь конечное число таких точек, в которых мера разрыва ее превосходит данное положительное число. Для функций

немонотонных дело обстоит иначе; так, мера разрыва функции Дирихле в каждой точке равна единице.

Наконец, если функция $f(x)$ монотонна на отрезке $[a, b]$, то пределы $f(c+0)$ и $f(c-0)$ в силу леммы 1′ лекции I (стр. 31) существуют в каждой точке c этого отрезка (в точке a существует лишь первый, в точке b — лишь второй из этих пределов); отсюда следует, что монотонная функция может иметь точки разрыва только первого рода.

Функции с ограниченной вариацией. Пусть на отрезке $[a, b]$ определена функция $f(x)$. Разобьем этот отрезок любым способом на любое число n частичных отрезков, обозначим точки деления слева направо через $x_1, x_2, \ldots, x_{n-1}$ и положим еще для общности $a = x_0$ и $b = x_n$, так что

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b;$$

сумма

$$S = \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$$

зависит от произведенного нами разбиения отрезка $[a, b]$; для различных разбиений эта сумма имеет, вообще говоря, различную величину; производя указанное разбиение всевозможными способами (причем и число n может меняться произвольным образом), мы получаем бесконечное множество сумм S. Верхнюю грань M этого множества (которая может, конечно, быть и равной $+\infty$) называют *полной вариацией* функции $f(x)$ на отрезке $[a, b]$; мы будем обозначать ее через $V_f(a, b)$. *Если* $V_f(a, b)$ *есть число* ($V_f(a, b) < +\infty$), *то функцию* $f(x)$ *называют функцией с ограниченной вариацией на отрезке* $[a, b]$. При $V_f(a, b) = +\infty$ говорят, что вариация функции $f(x)$ на отрезке $[a, b]$ бесконечна (или не ограничена).

Класс функций с ограниченной вариацией играет весьма заметную роль в анализе и его приложениях. Очевидно, в частности, что всякая функция, монотонная на отрезке $[a, b]$, есть функция с ограниченной вариацией на этом отрезке; в самом деле, для монотонной функции $f(x)$ сумма S при любом разбиении равна

$|f(b) - f(a)|$, вследствие чего и

$$V_f(a, b) = |f(b) - f(a)|.$$

Для исследования общих свойств функций с ограниченной вариацией имеет основное значение одна теорема, в силу которой исследование самых общих типов таких функций может быть целиком сведено к изучению монотонных функций; именно, оказывается, что *всякая функция с ограниченной вариацией есть разность двух неубывающих функций* (или, что то же, сумма двух монотонных функций, из которых одна — неубывающая, а другая — невозрастающая). В силу этого предложения все основные свойства монотонных функций, касающиеся числа и характера возможных точек разрыва, распространяются и на все функции с ограниченной вариацией.

Чтобы доказать эту основоположную теорему, условимся для краткости писать $V(x)$ вместо $V_f(a, x)$ и положим

$$\frac{1}{2}[V(x) + f(x)] = P(x), \quad \frac{1}{2}[V(x) - f(x)] = N(x);$$

отсюда

$$f(x) = P(x) - N(x),$$

и наша теорема будет доказана, если мы покажем, что функции $N(x)$ и $P(x)$ — неубывающие на отрезке $[a, b]$.

Пусть $a \leqslant x_1 < x_2 \leqslant b$; так как

$$2[P(x_2)-P(x_1)] = V(x_2)-V(x_1)+[f(x_2)-f(x_1)],$$

$$2[N(x_2)-N(x_1)] = V(x_2)-V(x_1)-[f(x_2)-f(x_1)],$$

то для нашей цели достаточно показать, что

$$V(x_2) - V(x_1) \geqslant |f(x_2) - f(x_1)|. \qquad (5)$$

Пусть ε — произвольное положительное число; по определению величины $V(x_1)$, как верхней грани сумм S, существует такое разбиение отрезка $[a, x_1]$, что соответствующая сумма $S = S(a, x_1)$ превосходит $V(x_1) - \varepsilon$; но сумма $S(a, x_1) + |f(x_2) - f(x_1)| = S(a, x_2)$ есть, очевидно, одна из сумм S для отрезка

$[a, x_2]$ и потому не превосходит верхней грани $V(x_2)$ таких сумм; таким образом, мы получаем

$$V(x_1) - \varepsilon + |f(x_2) - f(x_1)| < S(a, x_1) + \\ + |f(x_2) - f(x_1)| = S(a, x_2) \leqslant V(x_2),$$

откуда

$$V(x_2) - V(x_1) > |f(x_2) - f(x_1)| - \varepsilon;$$

так как ε произвольно мало, то отсюда следует (5), и наша теорема доказана.

ЛЕКЦИЯ IV

РЯДЫ

Сходимость и сумма ряда.— Критерий Коши.— Ряды с положительными членами.— Абсолютная и условная сходимость.— Бесконечные произведения.— Ряды функций.— Степенные ряды.

Сходимость и сумма ряда. Сегодняшнюю лекцию мы посвятим одному из важнейших инструментов математического анализа — бесконечным рядам. Как вам известно, перед лицом больших аналитических проблем бесконечный ряд представляется нам чисто техническим орудием исследования, очень полезным и удобным, но скромным по своему принципиальному значению. Тем не менее орудие это вполне заслуживает того, чтобы даже в нашем кратком курсе ему была посвящена особая лекция; и не столько потому, что многочисленными применениями его проникнуто все здание как самого анализа, так и почти всех опирающихся на него прикладных наук, сколько по той причине, что на сравнительно несложном материале, какой представляет нам собою теория рядов, типичные для всего анализа ходы мыслей, цепи представлений и образов и даже целые логические схемы выступают с особенной ясностью и рельефностью; хорошо известно, что учащемуся, который активно и прочно овладел теорией рядов, дальнейшее усвоение основных разделов анализа обычно не доставляет уже никаких затруднений.

Бесконечным числовым рядом называется выражение вида

$$u_1 + u_2 + \ldots + u_n + \ldots, \tag{1}$$

или

$$\sum_{n=1}^{\infty} u_n,$$

где все u_n — числа, которые мы будем предполагать вещественными. Числа u_n называются *членами* ряда. Конечные суммы

$$s_n = \sum_{k=1}^{n} u_k$$

мы будем называть *частичными суммами* ряда (1).

Основной вопрос, который ставится в отношении каждого данного ряда (1), есть вопрос о его *сходимости*. Если предел

$$\lim_{n\to\infty} s_n = s$$

существует, ряд (1) называется *сходящимся*, а число s — его *суммой*. В противном случае ряд (1) называется *расходящимся* и суммы не имеет. Случай, когда $s_n \to \infty$ или $s_n \to -\infty$ при $n \to \infty$, мы формально вправе, конечно, отнести к любой из этих двух разновидностей бесконечных рядов; обычно в этом случае признают ряд расходящимся, т. е. под суммой ряда всегда разумеют некоторое число; однако при этом мы, объединяя под одним именем расходящихся как ряды с бесконечными суммами, так и ряды, вовсе не имеющие сумм, должны помнить, что эти два типа рядов не имеют в своей основе ничего общего и объединяются лишь потому, что оба противостоят (хотя и в разных смыслах) рядам с конечными суммами.

В основных своих разделах анализ оперирует только сходящимися рядами. У того, кто впервые приступает к изучению рядов, первичное представление о сумме ряда и процессе ее получения неизбежно складывается по полной, часто не подвергаемой критическому анализу, аналогии с конечными суммами: s_n — это сумма n первых членов ряда, а s — сумма *всех* его членов. Очень часто и сам лектор стимулирует и даже прямо воспитывает такое представление. И надо признать, что если относиться к нему достаточно осторожно и не

утрачивать критической настроенности, это представление действительно способно как содействовать усвоению основных фактов теории рядов, так и эвристически предвидеть новые, еще не знакомые закономерности. Однако ни на одну минуту нельзя забывать и об опасности зайти в этой аналогии слишком далеко, а тем более — придавать ей доказательную силу. Как вы знаете, аналогия между бесконечными рядами и конечными суммами сколько-нибудь далеко может быть проведена только для так называемых абсолютно сходящихся рядов (о которых мы будем говорить ниже); для рядов же «условно сходящихся» наше стремление представить себе их суммы образующимися аналогично конечным суммам немедленно наталкивается на неодолимые препятствия; трудно, в самом деле, представлять себе «суммою *всех* членов ряда» такую сумму, величину которой мы можем изменить, изменив порядок слагаемых. Наконец, слишком крепко угнездившееся представление о сумме ряда, как о чем-то очень похожем на конечные суммы, таит в себе и серьезную методологическую опасность, угрожая выхолостить все специфическое идейное и предметное содержание суммы бесконечного ряда. Дело в том, что процесс образования суммы ряда вовсе не подобен процессу конечного суммирования и вовсе не состоит в том (как иногда себе представляют), что члены ряда прибавляются один за другим, «покуда не будут все исчерпаны»; это было бы, разумеется, совершенно безнадежным предприятием; исчерпать члены ряда последовательным прибавлением нельзя, ибо их бесконечное множество, и если мы все же имеем возможность говорить о сумме ряда, то это именно потому, что мы отказываемся от этого безнадежного процесса бесконечного прибавления и заменяем его совсем другой операцией (предельным переходом), сразу приводящей нас к цели. Вот подробное описание того, как формируется сумма бесконечного ряда; начинаем мы, разумеется, как и в случае конечных сумм, с последовательного прибавления, т. е. образуем суммы s_1, s_2 и т. д.; но мы вовсе не имеем в виду продолжать этот процесс безгранично; образуя частичные суммы, мы внимательно исследуем их при-

роду и строение, прежде всего — их зависимость от числа n взятых членов; другими словами, предметом нашего изучения служит *величина s_n как функция от n* (во многих случаях для получения полной картины этой функциональной зависимости вовсе нет надобности в изучении сумм s_1, s_2 и т. д. *небольшого* числа членов, а удается сразу получить, например, удобное аналитическое выражение функции s_n; именно так обстоит дело, например, в известном из элементарной алгебры случае геометрической прогрессии). Точнее, мы стараемся узнать, стремится ли величина s_n при $n \to \infty$ к некоторому пределу и если стремится, то к какому именно. Таким образом, формирование суммы бесконечного ряда слагается из двух последовательных моментов: 1) формирования частичных сумм s_n и изучения их зависимости от n; 2) перехода к пределу при $n \to \infty$. Вы, конечно, видите, что все это весьма мало похоже на «прибавление до полного исчерпания»; и забывать об этом различии, конечно, ни на минуту нельзя, если мы не хотим иметь над собой постоянной угрозы впадать в ошибки на базе необоснованных аналогий.

Сделаем еще следующее замечание. Мы видим, что вопрос о сумме данного ряда (1) целиком сводится к вопросу о пределе связанной с этим рядом последовательности

$$s_1, s_2, \ldots, s_n, \ldots; \tag{2}$$

обратно, если задана совершенно произвольная последовательность (2), мы всегда можем, полагая

$$u_1 = s_1, \quad u_2 = s_2 - s_1, \quad u_3 = s_3 - s_2, \quad \ldots, \quad u_n = s_n - s_{n-1}, \quad \ldots,$$

свести вопрос о пределе последовательности (2) к вопросу о сумме ряда (1); таким образом, с этой точки зрения основная задача теории рядов ничем не отличается от основной задачи теории пределов (для последовательностей).

На самом деле, конечно, специфика теории рядов, ее задач и ее методов обусловлена именно той особой концепцией, которую получают члены последовательности (2), когда мы рассматриваем их как последовательные частичные суммы некоторого ряда (1); и эта

концепция оказывается достаточно интересной в идейном и достаточно важной в прикладном отношении, чтобы составить собою предмет особой теории; эта теория и есть теория бесконечных рядов.

Критерий Коши. *Если ряд* (1) *сходится, то* $u_n \to 0$ *при* $n \to \infty$. В самом деле, $u_n = s_n - s_{n-1}$ при $n > 1$; так как при $n \to \infty$ s_n и s_{n-1} имеют один и тот же предел s, то $u_n \to 0$. Это свойство сходящихся рядов важно потому, что во многих случаях позволяет легко установить расходимость ряда; для этого достаточно показать, что u_n при $n \to \infty$ не является бесконечно малой величиной. Однако, как вы знаете, из того, что $u_n \to 0$ при $n \to \infty$, нельзя еще заключить, что ряд (1) сходится; так, «гармонический» ряд

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots,$$

как вам известно, расходится ($s_n \to \infty$ при $n \to +\infty$), хотя n-й член его и бесконечно мал при $n \to \infty$. Такая односторонность рассматриваемого признака (необходимость без достаточности), конечно, в значительной мере ограничивает сферу его применения.

Вопрос о сходимости ряда есть, как мы видели, только специальная форма вопроса о существовании предела у некоторой числовой последовательности. Поэтому мы получаем необходимый и достаточный признак сходимости ряда, переводя на язык теории рядов известный нам уже критерий Коши существования предела у последовательности (лекция II, стр. 47). Для того чтобы последовательность частичных сумм s_n ряда (1) имела предел при $n \to \infty$, необходимо и достаточно, как мы знаем, чтобы при любом $\varepsilon > 0$ для всех достаточно больших n и для любого $k \geqslant 0$ выполнялось неравенство

$$|s_{n+k} - s_n| < \varepsilon.$$

Так как $s_{n+k} - s_n = \sum\limits_{i=n+1}^{n+k} u_i$, то мы непосредственно приходим к следующему предложению.

Критерий Коши. *Для сходимости ряда* (1) *необходимо и достаточно, чтобы для любого* $\varepsilon > 0$ *при*

всех достаточно больших n и любом натуральном k выполнялось неравенство

$$|u_{n+1}+u_{n+2}+\ldots+u_{n+k}|<\varepsilon. \qquad (3)$$

Выражаясь образно, условие критерия Коши состоит в том, чтобы достаточно удаленный и как угодно длинный «кусок» ряда становился сколь угодно малым по абсолютному значению. Разумеется, как и в теории пределов, критерий Коши может только установить существование суммы ряда, но ничего не говорит нам о ее величине.

Критерий Коши, будучи очень плодотворным орудием общетеоретического исследования, сравнительно редко применяется, как мы уже говорили по аналогичному поводу в лекции II, к установлению сходимости индивидуальных, конкретных рядов; причиной этого является то, что обычно бывает нелегко установить, выполняется ли для данного конкретного ряда условие (3).

Поэтому, как вы знаете, теория рядов строит большое число других признаков сходимости; эти признаки, не имея присущей критерию Коши характеристичности (т. е. необходимости и достаточности одновременно), зато с несравненно бóльшим удобством применяются к отдельным конкретным рядам. Большинство этих признаков относится к рядам с положительными членами. Этот простейший и вместе с тем важнейший класс рядов методологически удобно рассмотреть в самом начале, чтобы затем постараться по возможности свести изучение произвольного ряда к исследованию рядов этого простейшего типа.

Ряды с положительными членами. Если все члены ряда (1) неотрицательны, то s_n, очевидно, есть неубывающая функция от n; в силу леммы 1 лекции I возможен поэтому лишь один из двух случаев: либо ряд (1) сходится, либо $\lim\limits_{n\to\infty} s_n=+\infty$; другими словами, ряд (1) сходится или расходится, смотря по тому, ограничена или не ограничена величина s_n при $n\to\infty$. Все признаки сходимости для рядов этого класса устанавливаются поэтому таким образом, что доказывается ограниченность функции s_n при соответствующих условиях.

В основе большинства этих признаков лежит следующий, чрезвычайно мощный принцип сравнения:

Если все члены рядов

$$u_1 + u_2 + \ldots + u_n + \ldots \qquad (A)$$

и

$$v_1 + v_2 + \ldots + v_n + \ldots \qquad (B)$$

неотрицательны и если при $n \geqslant n_0$

$$u_n \leqslant v_n, \qquad (C)$$

то из сходимости ряда (B) *вытекает сходимость ряда* (A) (и, значит, из расходимости ряда (A) вытекает расходимость ряда (B)).

Доказательство этого основного принципа очень просто: обозначим соответственно через s_n и s_n' частичные суммы рядов (A) и (B); если ряд (B) сходится, то величина s_n' ограничена; а так как в силу условия (C) при $n \geqslant n_0$

$$s_n - s_{n_0} \leqslant s_n' - s_{n_0}',$$

то величина $s_n \leqslant s_n' + (s_{n_0} - s_{n_0}')$ также ограничена, а следовательно, сходится и ряд (A).

Взяв в качестве ряда (B) какой-нибудь ряд, сходимость которого уже установлена, обычно затем доказывают, что в случае, когда ряд (A) подчиняется какому-либо определенному условию, члены его необходимо связываются с членами ряда (B) соотношением (C); всякое такое условие и может служить признаком сходимости для ряда (A). Так, если выбрать в качестве ряда (B) обыкновенную геометрическую прогрессию, то мы этим путем получаем для ряда (A) как раз простейшие, вероятно, хорошо известные вам признаки сходимости Даламбера, Коши и т. п. Как эти, так и все другие признаки всегда представляют собой условия, в той или другой конкретной форме высказывающие требование, чтобы члены ряда (A) *достаточно быстро убывали* с ростом n; так, известный признак Даламбера требует, чтобы отношение u_{n+1}/u_n при достаточно большом n оставалось не больше числа, меньшего единицы, что очевидным образом характеризует

собой именно достаточную быстроту убывания величины u_n с ростом n.

Здесь нам нет надобности останавливаться на признаках этого типа; вы найдете их, вместе с исчерпывающими доказательствами, в большом числе в каждом курсе анализа. Вместо этой, в основном формальной, теории постараемся лучше повнимательнее вглядеться в наши представления о *быстроте сходимости* ряда с положительными членами. Вопрос этот, хотя им почти никогда не интересуются курсы, предназначенные для будущих практических работников, как раз имеет, независимо от своего принципиального интереса, и непосредственно практическое значение. В самом деле, если ряд сходится медленно, т. е. если для получения суммы s_n, сколько-нибудь близкой к предельному значению s, приходится суммировать очень большое число n его членов, то такой ряд, при всей своей теоретической полноценности, не может служить инструментом для приближенного вычисления числа s и поэтому в практическом отношении, по крайней мере непосредственно, не может принести значительной пользы. Заметим вскользь, что иногда приходится встречаться и с обратным положением вещей: расходящийся ряд при известных условиях может оказаться очень удобным орудием для практического вычисления некоторых величин. Это обстоятельство дало повод известному французскому ученому Анри Пуанкаре, занимавшемуся исследованием этого явления, высказать парадоксальную мысль о том, что сходящиеся ряды иногда оказываются «практически расходящимися» и, наоборот, расходящиеся ряды — «практически сходящимися».

Для всякого сходящегося ряда «остаток» $r_n = s - s_n = u_{n+1} + u_{n+2} + \ldots$ есть величина бесконечно малая при $n \to \infty$; быстрота сходимости ряда целиком зависит от порядка этой бесконечной малой, т. е. от того, насколько быстро она стремится к нулю при $n \to \infty$. Обычно считают, что ряд сходится с хорошей быстротой, если при $n \to \infty$ остаток r_n убывает подобно члену геометрической прогрессии; для этого достаточно, чтобы члены u_n данного ряда при достаточно большом n сами не превосходили соответствующих членов

некоторой геометрической прогрессии; в самом деле, из

$$u_n < aq^n \qquad (n \geqslant n_0),$$

где $a > 0$ и q $(0 < q < 1)$ — постоянные, вытекает, что при $n \geqslant n_0$

$$r_n < a(q^{n+1} + q^{n+2} + \ldots) = \frac{aq}{1-q}q^n.$$

Значительно медленнее сходятся ряды, у которых с ростом n остаток убывает, как некоторая отрицательная степень числа n, т. е., например, имеет порядок $\frac{1}{n^2}$, $\frac{1}{n^3}$ и т. п. Такие ряды часто представляют уже затруднения для практического использования. При этом надо иметь в виду, что если u_n при достаточно большом n оказывается меньше, чем an^{-k}, где $a > 0$ и $k > 1$ — постоянные (причем число k отнюдь не обязательно целое), то $r_n < bn^{-k+1}$, где b — другая положительная постоянная; в самом деле, применяя к функции $f(x) = x^{k-1}$ теорему Лагранжа о конечном приращении, мы находим при $x > 0$

$$(x+1)^{k-1} - x^{k-1} = (k-1)(x+\theta)^{k-2} > (k-1)x^{k-2} \qquad (0 < \theta < 1);$$

отсюда, полагая $x = n - 1$ и деля обе части на $n^{k-1}(n-1)^{k-1}$,

$$(n-1)^{-k+1} - n^{-k+1} =$$

$$= \frac{n^{k-1} - (n-1)^{k-1}}{n^{k-1}(n-1)^{k-1}} > \frac{(k-1)(n-1)^{k-2}}{n^{k-1}(n-1)^{k-1}} > \frac{k-1}{n^k},$$

и, следовательно,

$$\frac{1}{n^k} < \frac{1}{k-1}\left[(n-1)^{-k+1} - n^{-k+1}\right];$$

если $u_n < an^{-k}$, то это дает

$$u_n < \frac{a}{k-1}\left[(n-1)^{-k+1} - n^{-k+1}\right]$$

и, следовательно,

$$r_n = \sum_{i=1}^{\infty} u_{n+i} < \frac{a}{k-1} \sum_{i=1}^{\infty} \left[(n+i-1)^{-k+1} - (n+i)^{-k+1}\right] = \frac{a}{k-1} n^{-k+1},$$

что мы и утверждали.

Возвращаясь теперь к вопросу о признаках сходимости, сделаем следующее замечание. Для всякого сходящегося ряда $u_n \to 0$ при $n \to \infty$; однако даже для ряда с положительными членами это стремление «общего члена» u_n к нулю может не быть монотонным; простым примером может служить ряд, полученный из геометрической прогрессии с помощью перестановки u_1 с u_2, u_3 с u_4 и т. д., т. е., например, ряд

$$\frac{1}{2^2} + \frac{1}{2} + \frac{1}{2^4} + \frac{1}{2^3} + \ldots + \frac{1}{2^{2n}} + \frac{1}{2^{2n-1}} + \ldots$$

Однако в подавляющем большинстве случаев конкретные числовые ряды с положительными членами, встречающиеся в анализе и его приложениях, обладают тем свойством, что $u_{n+1} \leqslant u_n$ $(n = 1, 2, \ldots)$, т. е. $u_n \to 0$ в этих рядах, монотонно убывая с возрастанием n. Поэтому заслуживает большого внимания группа специальных критериев сходимости рядов с монотонно убывающими положительными членами, тем более что в этой группе имеются признаки характеристические (т. е. необходимые и достаточные) и в то же время часто находящие себе удобное применение к исследованию отдельных конкретных рядов. Мы рассмотрим два из этого рода признаков.

Теорема. (Интегральный признак Коши.) *Пусть $f(x)$ — положительная, непрерывная и невозрастающая функция на полупрямой $0 \leqslant x < +\infty$. Тогда для сходимости ряда*

$$f(1) + f(2) + \ldots + f(n) \ldots \qquad (A)$$

необходимо и достаточно, чтобы существовал конечный предел при $a \to +\infty$ интеграла

$$\int_0^a f(u)\, du \qquad (B)$$

(для чего, в свою очередь, необходимо и достаточно, разумеется, чтобы этот интеграл оставался ограниченным при $0 \leqslant a < +\infty$).

Для удобства расположения доказательства мы формулировали этот признак, отправляясь от данной функции, а не от данного ряда; ясно, однако, что для любого ряда (A) с монотонно убывающими положительными членами можно легко построить сколько угодно функций $f(x)$, удовлетворяющих требованиям нашей теоремы и таких, что $f(n) = u_n$ $(n \geqslant 1)$; таким образом, наша теорема действительно служит характеристическим критерием сходимости для рядов указанного типа.

Доказательство. Так как, очевидно, в силу монотонности (невозрастания) функции $f(x)$

$$\int_{n-1}^{n} f(u)\,du \geqslant f(n) \geqslant \int_{n}^{n+1} f(u)\,du,$$

то

$$\int_{0}^{n} f(u)\,du \geqslant \sum_{k=1}^{n} f(k) \geqslant \int_{1}^{n+1} f(u)\,du;$$

эти неравенства непосредственно показывают, что при $n \to \infty$ величины $\sum_{k=1}^{n} f(k)$ и $\int_{0}^{n} f(u)\,du$ либо обе ограничены, либо обе не ограничены, чем теорема и доказана, так как в обоих случаях ограниченность равносильна существованию конечного предела.

Применения этого критерия к отдельным рядам, как мы уже заметили, во многих случаях очень просты. Пусть, например, $u_n = n^{-\alpha}$, где $\alpha > 0$. Полагая

$$f(x) = \begin{cases} 1 & (0 \leqslant x \leqslant 1), \\ x^{-\alpha} & (1 \leqslant x < +\infty), \end{cases}$$

мы, очевидно, получаем функцию, удовлетворяющую всем требованиям нашей теоремы; но при $\alpha \neq 1$ и $a > 1$

$$\int_{0}^{a} f(x)\,dx = 1 + \int_{1}^{a} x^{-\alpha}dx = 1 + \frac{a^{1-\alpha} - 1}{1-\alpha},$$

откуда в силу интегрального признака наш ряд сходится при $\alpha > 1$ и расходится при $\alpha < 1$; наконец, при $\alpha = 1$

$$\int_0^a f(x)\,dx = 1 + \int_1^a \frac{dx}{x} = 1 + \ln a \to \infty$$

при $a \to +\infty$, т. е. гармонический ряд $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ расходится.

Переходим ко второму критерию.

Теорема. *Если $u_n \geqslant u_{n+1} > 0$, то для сходимости ряда* (1) *необходимо и достаточно, чтобы сходился ряд*

$$u_1 + 2u_2 + 4u_4 + 8u_8 + \ldots + 2^n u_{2^n} + \ldots \qquad (4)$$

Доказательство. В силу монотонного убывания членов u_n ряда (1)

$$2^{n-1} u_{2^{n-1}} \geqslant \sum_{k=2^{n-1}+1}^{2^n} u_k \geqslant 2^{n-1} u_{2^n}$$

(число членов суммы в средней части неравенств есть 2^{n-1}), откуда

$$\sum_{k=1}^{n-1} 2^k u_{2^k} \geqslant \sum_{k=1}^{2^n} u_k \geqslant \frac{1}{2} \sum_{k=1}^{n} 2^k u_{2^k}.$$

Эти неравенства непосредственно показывают, что частичные суммы рядов (1) и (4) либо одновременно ограничены, либо одновременно не ограничены, откуда и следует утверждение теоремы.

В случае $u_n = n^{-\alpha}$ $(\alpha > 0)$ ряд (4) имеет общим членом $2^{n(1-\alpha)}$; в силу последней теоремы ряд $\sum\limits_{n=1}^{\infty} n^{-\alpha}$ поэтому сходится при $\alpha > 1$ и расходится при $\alpha \leqslant 1$.

Абсолютная и условная сходимость. Мы переходим теперь к рядам с членами произвольного знака; как вы знаете, сходящиеся ряды этого общего типа принято с самого начала разделять на два класса, действительно существенно различных по своим свойствам: ряды, аб-

солютно сходящиеся, и ряды, сходящиеся условно. Мы говорим, что ряд (1) сходится *абсолютно*, если сходится ряд

$$|u_1| + |u_2| + \ldots + |u_n| + \ldots; \qquad (5)$$

если же ряд (1) сходится, а ряд (5) расходится, то мы называем ряд (1) *условно* сходящимся.

Здесь прежде всего надо обратить внимание на то, что *абсолютная* сходимость данного ряда определяется как сходимость некоторого *другого* ряда; поэтому совсем нетривиальной является (известная вам, конечно) теорема: *всякий абсолютно сходящийся ряд сходится.*

Смысл этой теоремы состоит, конечно, в том, что из сходимости ряда (5) вытекает сходимость ряда (1); доказательство ее проще всего проводится на основе критерия Коши: так как при любых $n > 0$, $k > 0$

$$\left|\sum_{i=1}^{k} u_{n+i}\right| \leqslant \sum_{i=1}^{k} |u_{n+i}|$$

и так как, согласно критерию Коши, в случае сходимости ряда (5) правая часть сколь угодно мала при всех достаточно больших n и любом k, то то же справедливо и для левой части, откуда, в силу того же критерия Коши, вытекает сходимость ряда (1).

Абсолютно сходящиеся ряды обнаруживают в своих свойствах далеко идущее сходство с конечными суммами: их можно почленно перемножать, как конечные суммы (распределительный закон); члены их можно произвольно перемещать (переместительный закон) и группировать (сочетательный закон), не нарушая этим сходимости ряда и не изменяя его суммы (в частности, всеми этими свойствами обладают, конечно, ряды с положительными членами, сходимость которых всегда имеет абсолютный характер). Мы не будем здесь останавливаться на доказательствах этих свойств; они чисто формальны, и те изложения, которые вы найдете в руководствах, не вызовут у вас никаких затруднений.

Вглядимся зато несколько подробнее в условно сходящиеся ряды; вот перед нами *сходящийся* ряд (1);

представьте себе, например, известный ряд Лейбница

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots;$$

все обстоит благополучно, частичные суммы стремятся к пределу, остаток стремится к нулю (и даже не слишком медленно), как у всякого сходящегося ряда. Но как шатко это благополучие! Стоит только надлежащим образом изменить порядок членов, и ряд получает другую сумму, а то и вовсе перестает сходиться. И все это потому, что ряд (5) (в нашем примере — это гармонический ряд) расходится. Вы понимаете, конечно, что такое суммирование, при котором результат зависит от порядка слагаемых, не может приводить к конструкциям, сколько-нибудь аналогичным конечным суммам; самое наименование «условно сходящиеся ряды» имеет, по-видимому, своей психологической предпосылкой желание называть такие ряды сходящимися только с известной оговоркой.

Что касается возможности изменять сумму и даже нарушать сходимость условно сходящихся рядов с помощью перестановки их членов, то в этом направлении любой условно сходящийся ряд действительно предоставляет нам совершенно неограниченные возможности, как показывает следующая простая, но изящная

Теорема. *Если ряд* (1) *сходится условно и если* α *— любое число или любой из символов* $\pm\infty$, *то надлежаще выбранной перестановкой членов этого ряда всегда можно заставить новый ряд сходиться к* α.

Для доказательства установим прежде всего одну лемму, которая и сама по себе не лишена интереса: *если ряд* (1) *сходится условно, то его неотрицательные члены образуют расходящийся ряд* (1′), *и точно так же его отрицательные члены образуют расходящийся ряд* (1″). В самом деле, пусть $s_n = s_n' + s_n''$, где s_n' — сумма неотрицательных, а s_n'' — отрицательных членов ряда (1), входящих в частичную сумму s_n. Если бы оба ряда (1′) и (1″) сходились, то s_n' и s_n'' имели бы пределы при $n \to \infty$, а следовательно, и $s_n' - s_n'' = |u_1| + |u_2| + \ldots + |u_n|$ стремилось бы к некоторому пределу, т. е. ряд (1) был бы абсолютно сходящимся; если бы один из

рядов (1′) и (1″), например (1′), был расходящимся, а другой, (1″), сходящимся, то при $n \to \infty$ мы имели бы $s_n' \to +\infty$, $s_n'' \to c$, где c — число, откуда $s_n = s_n' + s_n'' \to +\infty$, т. е. ряд (1) был бы расходящимся; таким образом, ряды (1′) и (1″) необходимо должны оба расходиться.

Переходим теперь к доказательству теоремы. Рассмотрим два случая:

1) α есть число; пусть для определенности $\alpha > 0$; будем тогда располагать члены ряда (1) в следующем порядке: сначала будем брать неотрицательные члены в их естественном порядке (т. е. в том порядке, в каком они встречаются в ряде (1)); при этом сумма взятых членов будет возрастать; как только эта сумма превзойдет α, мы назовем ее *поворотной* суммой и начнем присоединять к ней последовательно отрицательные члены ряда (1), опять-таки в их естественном порядке; при этом сумма взятых членов будет уменьшаться; как только она сделается меньше α, мы снова назовем полученную сумму *поворотной* и опять начнем ее увеличивать, прибавляя следующие по порядку неотрицательные члены ряда (1), и т. д. Продолжая этот процесс неограниченно, мы тем самым, очевидно, расположим члены ряда (1) в некотором определенном порядке. Мы должны будем показать, что вновь образованный ряд имеет сумму α; однако прежде этого мы должны еще осветить одну из деталей произведенной конструкции: откуда мы знаем, что мы можем, беря достаточно много неотрицательных членов ряда (1), получить сумму, большую чем α? Здесь приходит на помощь наша лемма: в силу нее ряд, составленный из неотрицательных членов ряда (1), расходится, и, следовательно, беря достаточно большое число этих неотрицательных членов, мы можем получить сумму как угодно большую — в частности, бóльшую чем α. В силу этой же леммы мы и в дальнейшем ходе нашей конструкции всякий раз можем быть уверены, что прибавление неотрицательных членов рано или поздно сделает сумму всех взятых членов больше α, а прибавление отрицательных — меньше α.

Убедившись таким образом в возможности нашей конструкции, мы должны теперь показать, что частичные суммы нового ряда, становясь то больше, то меньше чем α, безгранично приближаются к этому числу. Из нашего построения и из того, что $u_n \to 0$ при $n \to \infty$, очевидно, что если мы возьмем две последовательные поворотные суммы нового ряда, то все достаточно далекие его суммы будут уже заключены между ними. Поэтому нам достаточно доказать, что поворотные суммы будут иметь своим пределом α. Но ясно, что каждая поворотная сумма отстоит от α не больше, чем на абсолютную величину своего последнего члена, т. е. величину, заведомо стремящуюся к нулю, так как этот член есть член нашего первоначального (сходящегося) ряда. Таким образом, в этом случае теорема доказана.

2) Пусть теперь $\alpha = +\infty$ (в случае $\alpha = -\infty$ доказательство совершенно аналогично); так как $u_n \to 0$ при $n \to \infty$, то множество чисел $|u_n|$ $(n = 1, 2, \ldots)$ ограничено; пусть β — его верхняя грань; рассмотрим последовательность неотрицательных членов ряда (1) и разобьем ее на участки так, чтобы сумма членов каждого участка была больше, чем 2β (это возможно в силу доказанной нами леммы); расположим теперь члены ряда (1) так: сперва возьмем первый из только что построенных участков (в его естественном порядке), затем присоединим первый отрицательный член, затем второй участок, второй отрицательный член и т. д. Так как сумма членов каждого участка $> 2\beta$, а каждый отрицательный член по абсолютной величине $\leqslant \beta$, то после взятия n неотрицательных участков и n отрицательных членов мы получим сумму, превосходящую $2\beta n - \beta n = \beta n$, откуда и видно, что частичные суммы вновь построенного ряда безгранично возрастают при $n \to \infty$.

Мы нарочно с такой подробностью остановились на этих конструкциях; они столь же интересны, как и поучительны, заставляя нас вплотную вглядываться в структуру ряда, как бы логически анатомировать его. Для вас будет интересным и полезным упражнением показать, что надлежащей перестановкой членов из

каждого условно сходящегося ряда может быть получен ряд, не имеющий ни конечной, ни бесконечной суммы.

Бесконечные произведения. Наряду со сложением и другое арифметическое действие — умножение — может быть применено к столь угодно большому числу чисел; как там мы искали предел суммы безгранично возрастающего числа слагаемых, так и здесь мы можем ставить вопрос о пределе произведения безгранично возрастающего числа сомножителей; мы можем, подобно теории бесконечных рядов, создать теорию *бесконечных произведений.* Эта теория, созданная в основном уже очень давно, имеет, правда, до сих пор не столь исключительное число применений, как теория рядов; тем не менее область этих применений уже настолько широка (и с каждым десятилетием все более расширяется), что в наши дни она необходимо входит в арсенал орудий каждого образованного математика.

Бесконечным произведением называется выражение вида

$$z_1 z_2 \ldots z_n \ldots = \prod_{n=1}^{\infty} z_n; \qquad (6)$$

величины $\pi_n = z_1 z_2 \ldots z_n$ $(n = 1, 2, \ldots)$ мы называем *частичными произведениями*, а $\lim\limits_{n \to \infty} \pi_n$, если он существует,— величиной произведения (6).

Во всем дальнейшем мы будем считать всегда все сомножители z_n *положительными.* Такое ограничение, всегда принимаемое в области вещественных множителей, есть следствие отнюдь не желания упростить исследование, а просто того факта, что изучение произведений с отрицательными множителями, с одной стороны, изменением знака всего произведения сразу приводится к случаю положительных сомножителей, а с другой — не представляет реального интереса; к последнему заключению приводит целый ряд соображений; так, подобно тому свойству сходящихся рядов, что n-й член ряда стремится к нулю при $n \to \infty$, в сходящемся бесконечном произведении n-й множитель z_n должен (как легко предвидеть и как мы скоро дока-

жем) стремиться к единице, и, значит, все множители, начиная с некоторого из них, должны быть не только положительны, но даже сколь угодно близки к единице.

В случае положительных сомножителей исследование бесконечного произведения, очевидно, простым логарифмированием приводится к исследованию бесконечного ряда; полагая $v_n = \ln z_n, z_n = e^{v_n}$, мы видим, что сходимость произведения (6) равносильна сходимости ряда

$$v_1 + v_2 + \ldots + v_n + \ldots \tag{7}$$

и что сумма s этого ряда равна (в случае сходимости) логарифму величины π произведения (6). Есть только одно исключение из этого правила: если $\pi = 0$, то ряд (7), очевидно, расходится (частичные суммы имеют пределом $-\infty$); но бесконечные произведения, величина которых равна нулю, и во многих других отношениях имеют совершенно особое поведение, приближающее их скорее к расходящимся рядам, чем к сходящимся; так, в случае $\pi = 0$ не обязательно, чтобы было $z_n \to 1$ при $n \to \infty$ (пример: $z_n = \frac{1}{2}$, $n = 1, 2, \ldots$). Все эти обстоятельства привели к тому, что при $\pi = 0$ произведение (6) обычно называют расходящимся, так что сходимость этого произведения определяется как существование *положительного* предела $\pi = \lim\limits_{n\to\infty} \pi_n$. При таком определении сходимость произведения (6), очевидно, в точности эквивалентна сходимости ряда (7).

Возможность указанной нами полной редукции теории бесконечных произведений к исследованию рядов могла бы подать повод думать, что в построении особого учения о бесконечных произведениях нет никакой надобности. Однако, как и во многих подобных случаях, такое заключение было бы ошибочным; как это часто бывает в математике, новая *форма постановки* прежнего круга проблем здесь, с одной стороны, приводит к совершенно новой проблематике, а с другой — эвристически стимулирует методы решения поставленных задач.

В теории бесконечных произведений принято писать множители z_n в виде $1+u_n$, где, значит, $-1<u_n<+\infty$; при этом $\pi_n=\prod_{k=1}^{n}(1+u_k)$, а произведение (6) записывается в виде

$$(1+u_1)(1+u_2)\ldots(1+u_n)\ldots=\prod_{n=1}^{\infty}(1+u_n). \quad (8)$$

Если произведение сходится, то $\pi_n \to \pi$ при $n\to\infty$, где $0<\pi<+\infty$, вследствие чего при $n\to\infty$

$$u_n=\frac{\pi_n}{\pi_{n-1}}-1\to 0, \quad z_n=\frac{\pi_n}{\pi_{n-1}}\to 1.$$

Сопоставление произведения (8) с рядом (7) показывает, что рядам с положительными членами соответствуют произведения, у которых $z_n\geqslant 1$ или, что то же, $u_n\geqslant 0$ $(n=1, 2, \ldots)$ (а рядам с отрицательными членами — произведения, имеющие $z_n\leqslant 1$, $u_n\leqslant 0$). Это заставляет нас в первую очередь исследовать бесконечные произведения, у которых все u_n имеют один и тот же знак. Основное, наиболее употребительное предложение, касающееся таких произведений, составляет

Теорема. *Если все u_n имеют один и тот же знак, то для сходимости произведения* (8) *необходимо и достаточно, чтобы сходился ряд*

$$u_1+u_2+\ldots+u_n+\ldots \quad (9)$$

Доказательство. Прежде всего, мы вправе предположить, что $u_n\to 0$ при $n\to\infty$, так как в противном случае, как мы знаем, расходится как произведение (8), так и ряд (9), и, следовательно, теорема оказывается верной. В дальнейшем мы различаем два случая [1]).

---

[1]) Теорему можно доказать проще. Поскольку $\frac{\ln(1+|u_n|)}{|u_n|}\to 1$ при $n\to\infty$, найдется такой номер n_0, что при $n\geqslant n_0$ будет $\frac{1}{2}|u_n|<\ln(1+|u_n|)<2|u_n|$. Поэтому по признаку сравнения (см. стр. 99) ряд (9), имеющий члены одного знака, сходится одновременно с рядом (7), а значит, и с произведением (6). — *Прим. ред.*

1) Пусть $u_n \geqslant 0$ $(n = 1, 2, \ldots)$; так как при $x \to 0$ функция e^x отличается от $1 + x$ на положительную бесконечно малую второго порядка относительно x, то при достаточно большом n

$$e^{\frac{1}{2}u_n} \leqslant 1 + u_n \leqslant e^{u_n}; \tag{10}$$

обозначая через s_n частичные суммы ряда (9) и предполагая для простоты, что неравенства (10) выполняются для всех n (это не ограничивает общности рассуждения, так как в вопросах сходимости можно всегда отбросить, без влияния на результат, любое число начальных членов как ряда, так и произведения), мы находим почленным перемножением этих неравенств при $n = 1, 2, \ldots, m$

$$e^{\frac{1}{2}s_m} \leqslant \pi_m \leqslant e^{s_m};$$

если ряд (9) сходится, то $s_m \leqslant s < +\infty$, а потому $\pi_m < e^s < +\infty$, т. е. величина π_m остается ограниченной, и, следовательно, произведение (8) сходится. Обратно, если сходится произведение (8), то $\pi_m \leqslant \pi < < +\infty$, а значит,

$$\frac{1}{2}\, s_m \leqslant \ln \pi < +\infty,$$

т. е. сумма s_m остается ограниченной и ряд (10) сходится.

2) Пусть $u_n \leqslant 0$ $(n = 1, 2, \ldots)$; подобно предыдущему, находим (здесь $s_m \leqslant 0$)

$$e^{2s_m} \leqslant \pi_m \leqslant e^{s_m},$$

откуда, аналогично предыдущему, убеждаемся, что при $s_m \geqslant s > -\infty$ (случай сходимости ряда (9)) $\pi_m \geqslant e^{2s} > 0$, и произведение (8) сходится; обратно, при $\pi_m \geqslant \pi > 0$ мы имеем $s_m \geqslant \ln \pi_m \geqslant \ln \pi > -\infty$, т. е. из сходимости произведения (8) вытекает сходимость ряда (9).

Сопоставление произведения (6) или (8) с рядом (7) непосредственно приводит к *критерию Коши* для бесконечных произведений: *для сходимости произведе-*

ния (8) *необходимо и достаточно, чтобы, как бы мало ни было* $\varepsilon > 0$, *при всех достаточно больших* n *и любом* $k > 0$ *мы имели*

$$1 - \varepsilon < \prod_{i=n+1}^{n+k} (1 + u_i) < 1 + \varepsilon. \tag{11}$$

Однако для произведений, в которых все u_n сохраняют постоянный знак, этот критерий, в силу доказанной нами теоремы, заменяется другим, в большинстве случаев значительно более удобным (так как оценивать суммы бывает, вообще говоря, значительно легче, чем произведения); именно, мы можем, очевидно, ничего не меняя в словесной формулировке только что приведенного критерия, заменить неравенства (11) неравенством

$$\left| \sum_{i=n+1}^{n+k} u_i \right| < \varepsilon;$$

необходимо только помнить, что эта форма критерия Коши имеет силу лишь в случаях, когда все u_n сохраняют одинаковый знак.

К сожалению, в рамках настоящего курса мы по интересному вопросу о бесконечных произведениях должны ограничиться сказанным и перейти к следующей теме.

Ряды функций. До сих пор члены рассматриваемых нами рядов были числами. Но вы знаете, конечно, что анализу в первую очередь нужны *функциональные* ряды, т. е. такие ряды, члены которых являются функциями одной или нескольких независимых переменных. Для теории этих рядов все, что мы до сих пор рассматривали, играет роль подготовительного материала.

Пусть

$$u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots \tag{12}$$

есть ряд, члены которого — функции одной и той же независимой переменной x, определенные на некотором отрезке $[a, b]$. Давая этой переменной какое-нибудь числовое значение $x = \alpha$, мы превращаем ряд (12) в

обыкновенный числовой ряд $\sum_{n=1}^{\infty} u_n(\alpha)$; с точки зрения теории числовых рядов можно поэтому сказать, что формула (12) выражает собой не один ряд, а целый континуум различных рядов.

Ясно, конечно, что для функциональных рядов проблема сходимости ставится совсем иначе, чем для числовых; по отношению к ряду (12) не имеет смысла ставить вопрос о том, сходится он или расходится, так как, вообще говоря, он будет сходиться при одних значениях x и расходиться при других значениях. Осмысленная постановка вопроса здесь гласит: при каких значениях величины x на данном отрезке $[a, b]$ ряд (12) сходится и при каких расходится? Условимся называть *областью сходимости* данного ряда множество тех значений x, при которых он сходится, и *областью расходимости* — совокупность значений, делающих его расходящимся. Мы видим, что проблема сходимости для функционального ряда состоит в первую очередь в отыскании его области сходимости. Мы не будем здесь тратить времени на рассмотрение примеров, так как у нас их достаточно будет в дальнейшем.

Для всех аналитических применений теории функциональных рядов имеет основное значение понятие равномерной сходимости. Вот как лучше всего подойти к определению этого понятия. Если ряд (12) сходится в каждой точке некоторого множества M (т. е. для всех значений $x \in M$), то остаток

$$r_n(x) = s(x) - s_n(x)$$

этого ряда (будучи, конечно, подобно сумме ряда $s(x)$ и частичным суммам $s_n(x)$, функцией от x) для любого $x \in M$ стремится к нулю при $n \to \infty$. Для нашей цели нам потребуется более детально описать этот факт: каково бы ни было $x \in M$ и каково бы ни было $\varepsilon > 0$, существует такое n_0 (зависящее, конечно, как от ε, так и от x), что $|r_n(x)| < \varepsilon$ для всех $n \geqslant n_0$. Оставим теперь ε неизменным, но заставим число x пробегать все множество M; при этом для каждого x найдется свое n_0, т. е. свое «место» в данном ряде, начиная с ко-

торого $|r_n(x)| < \varepsilon$; но существует ли в этом ряде такое «место», начиная с которого неравенство $|r(x)| < \varepsilon$ выполнялось бы уже для *всех* $x \in M$? Очевидно, это зависит от того, будет ли множество тех чисел n_0, которые мы выше построили для различных x, ограниченным или нет; если среди этих чисел есть наибольшее, то оно, очевидно, и сможет служить тем «местом», начиная с которого $|r_n(x)| < \varepsilon$ для *всех* $x \in M$. Если же среди построенных нами чисел n_0 встречаются сколь угодно большие, то это показывает, что, как бы далеко мы ни продвигались в нашем ряде, всегда найдутся такие $x \in M$, для которых искомое «место» еще не достигнуто и лежит дальше; в этом случае выбор такого n_0, которое в указанном смысле годилось бы для всех $x \in M$, очевидно, невозможен.

Приведем пример этого последнего типа. Пусть для $n \geqslant 1$

$$s_n(x) = x^n(1 - x^n) \qquad (0 \leqslant x \leqslant 1),$$

т. е.

$$u_n(x) = x^n(1 - x^n) - x^{n-1}(1 - x^{n-1}).$$

Очевидно, $s_n(x) \to 0$ для любого $x \in [0, 1]$ при $n \to \infty$, так что

$$r_n(x) = -s_n(x) = -x^n(1 - x^n);$$

отсюда

$$r_n(2^{-1/n}) = -\frac{1}{4} \qquad (n = 1, 2, \ldots);$$

так как число $2^{-1/n}$ при любом n принадлежит отрезку $[0, 1]$, то, сколь бы велико ни было число n, неравенство $|r_n(x)| < \frac{1}{4}$ не может быть выполнено для всех $x \in [0, 1]$. Значит, мы имеем здесь, действительно, дело со вторым из вышеописанных случаев. На рис. 16 изображены для наглядности кривые $y = |r_1(x)|$ и (схематически) $y = |r_n(x)|$ для большого n. Кривая $y = |r_n(x)|$ при $x = 2^{-1/n}$ имеет максимум, равный $\frac{1}{4}$; влево же от этой точки всюду, за исключением ближайшего ее соседства, $|r_n(x)|$ чрезвычайно мало; при возрастании n точка $2^{-1/n}$ неограниченно прибли-

жается к 1. Из этого рисунка мы непосредственно видим, как протекает явление; тот парадоксальный на первый взгляд факт, что для каждого x мы имеем $r_n(x) \to 0$ при $n \to \infty$, в то время как для каждого n существует такое x, что $|r_n(x)| = \frac{1}{4}$, объясняется именно тем, что точка, в которой $|r_n(x)|$ достигает этого максимального значения, при возрастании n не остается на месте, а продвигается вправо, безгранично приближаясь к 1 при $n \to \infty$; это явление «отставания» $r_n(x)$ в его стремлении к нулю перемещается, таким образом, все ближе к точке 1, когда n возрастает; мы невольно ожидаем, что в самой точке 1 это отставание, сгустившись, создаст расходимость ряда; но на самом деле в точке 1 все обстоит в высшей степени благополучно, так как здесь $r_n(x) = 0$ при любом n.

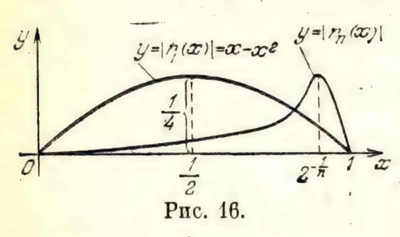


Рис. 16.

Вернемся теперь к тем двум случаям, которые мы условились различать выше. В первом из них мы говорим, что ряд (12) сходится равномерно на множестве M, а во втором — неравномерно. Таким образом, *ряд* (12) *равномерно сходится на некотором множестве* M, *если при сколь угодно малом* $\varepsilon > 0$ *существует такое* n_0 (*зависящее только от* ε), *что для всех* $n \geqslant n_0$ *и для всех* $x \in M$

$$|r_n(x)| < \varepsilon.$$

Мы подробно рассмотрели выше пример, показывающий типичную картину неравномерной сходимости. Заметим еще, что в точности таким же образом определяется равномерная сходимость *последовательности* функций

$$f_1(x),\ f_2(x),\ \ldots,\ f_n(x),\ \ldots$$

к предельной функции $f(x)$, если под $r_n(x)$ подразумевать разность $f(x) - f_n(x)$; мы можем, таким образом, сказать, что равномерная сходимость ряда равносиль-

на равномерной сходимости последовательности его частичных сумм.

Если для арифметического оперирования над рядом важна, как мы видели, его *абсолютная* сходимость, то для значительной части умозаключений аналитического характера очень удобной, а иногда и неизбежной предпосылкой служит *равномерная* сходимость его; мы сейчас рассмотрим несколько примеров, подтверждающих это.

В первую очередь рассмотрим вопрос о том, при каких условиях можно быть уверенным в непрерывности сумм $s(x)$ ряда (12), если известно, что все его члены непрерывны на отрезке $[a, b]$. Что сумма $s(x)$ может при этом оказаться и разрывной, показывает следующий пример: $u_n(x) = x^{n-1} - x^n$ $(n \geqslant 1,\ 0 \leqslant x \leqslant 1)$; $s_n(x) = 1 - x^n$;

$$s(x) = \begin{cases} 1 & \text{при} \quad 0 \leqslant x < 1, \\ 0 & \text{при} \quad x = 1. \end{cases}$$

Покажем теперь, что *если ряд* (12) *сходится равномерно, то из непрерывности его членов вытекает непрерывность суммы.* В самом деле, пусть на отрезке $[a, b]$ ряд (12) сходится равномерно и пусть все $u_n(x)$ (а значит, и все частичные суммы $s_n(x)$) непрерывны на $[a, b]$; пусть α — любое число отрезка $[a, b]$ и ε — сколь угодно малое положительное число; в силу предположенной равномерной сходимости ряда (12) мы можем взять n столь большим, что

$$|s_n(x) - s(x)| < \frac{\varepsilon}{3}$$

для всех $x \in [a, b]$. Закрепив временно это значение n и пользуясь тем, что функция $s_n(x)$ непрерывна на отрезке $[a, b]$ и значит, в частности, в точке α, мы можем утверждать, что для всех точек x некоторой окрестности U точки α

$$|s_n(x) - s_n(\alpha)| < \frac{\varepsilon}{3};$$

но в таком случае для любого $x \in U$

$$|s(x) - s(\alpha)| \leqslant$$
$$\leqslant |s(x) - s_n(x)| + |s_n(x) - s_n(\alpha)| + |s_n(\alpha) - s(\alpha)| <$$
$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

а это и показывает, что функция $s(x)$ непрерывна при $x = \alpha$.

Предложение, обратное только что доказанному, было бы неверным. Сумма ряда может оказаться в некоторых случаях непрерывной функцией и при неравномерной сходимости этого ряда; это можно усмотреть из примера, приведенного нами на стр. 115, где ряд сходится неравномерно, между тем как сумма его равна нулю на всем рассматриваемом отрезке. Равномерная сходимость ряда непрерывных функций, гарантируя непрерывность его суммы, не является, таким образом, *необходимой* предпосылкой этой непрерывности; теория, несколько более развитая, чем приведенные нами соображения, позволяет установить и характер сходимости, служащий необходимой и достаточной предпосылкой непрерывности суммы; однако на практике почти всегда непрерывность суммы устанавливается как следствие равномерной сходимости ряда.

Другой вопрос, в котором идея равномерной сходимости находит себе существенное применение,— это задача о «почленной интеграции» функционального ряда. Пусть все члены ряда (12) — функции непрерывные (и, следовательно, интегрируемые) на отрезке $[a, b]$; можно ли утверждать, что в таком случае и сумма $s(x)$ этого ряда интегрируема на этом отрезке и что

$$\int_a^b s(x)\,dx = \sum_{n=1}^{\infty} \int_a^b u_n(x)\,dx, \tag{13}$$

подобно тому, как мы это имеем для конечных сумм? Легко видеть, что во всех рассмотренных нами до сих пор примерах равенство (13) действительно имеет

место. Оно равносильно, конечно, любому из равенств

$$\lim_{n\to\infty}\int_a^b s_n(x)\,dx=\int_a^b s(x)\,dx,$$

$$\lim_{n\to\infty}\int_a^b r_n(x)\,dx=0.$$

Можно, однако, показать, что эти соотношения не всегда имеют место. Прежде всего, может оказаться,

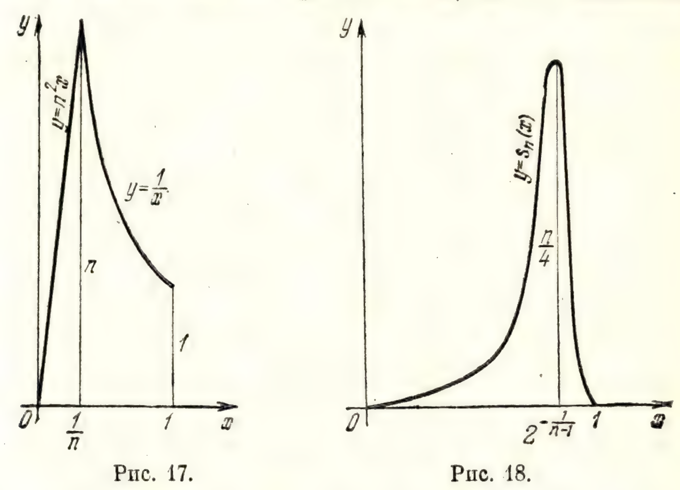


Рис. 17. Рис. 18.

что функция $s(x)$ неинтегрируема на отрезке $[a, b]$; для этого достаточно, например, выбрать (рис. 17)

$$s_n(x)=\begin{cases} n^2x & \left(0\leqslant x\leqslant\frac{1}{n}\right),\\ \frac{1}{x} & \left(\frac{1}{n}\leqslant x\leqslant 1\right)\end{cases}$$

(соответственное выражение для $u_n(x)=s_n(x)-s_{n-1}(x)$ вы легко установите сами); функция

$$s(x)=\begin{cases}\frac{1}{x} & \text{при}\quad 0<x\leqslant 1,\\ 0 & \text{при}\quad x=0,\end{cases}$$

как вы знаете, неинтегрируема на отрезке [0, 1] (ибо $\int_a^1 s(x)\,dx = \ln\frac{1}{a} \to \infty$ при $a \to 0$).

Но может случиться и так, что функция $s(x)$ интегрируема на отрезке $[a, b]$, а равенство (13) все же оказывается неверным. Например, пусть (рис. 18)

$$s_n(x) = nx^{n-1}(1 - x^{n-1}) \qquad (n \geqslant 1,\ 0 \leqslant x \leqslant 1).$$

Так как $\lim\limits_{n\to\infty} nx^{n-1} = 0$ при $0 \leqslant x < 1$[1]), то $s(x) = 0$ при $0 \leqslant x \leqslant 1$ и, значит,

$$\int_0^1 s(x)\,dx = 0;$$

между тем, как легко вычислить,

$$\int_0^1 s_n(x)\,dx = n\int_0^1 (x^{n-1} - x^{2n-2})\,dx = \frac{1}{2} - \frac{1}{2(2n-1)},$$

вследствие чего равенство (13) действительно не имеет места. Но легко убедиться, что *для всякого равномерно сходящегося ряда непрерывных функций равенство* (13) *справедливо.* Прежде всего, в этом случае, как мы уже знаем, функция $s(x)$ непрерывна и, следовательно, интегрируема на отрезке $[a, b]$; далее, для любого $\varepsilon > 0$ существует такое целое число n_0, что при $n \geqslant n_0$ и любом x из отрезка $[a, b]$

$$|r_n(x)| < \varepsilon.$$

Применяя известную вам теорему интегрального исчисления, согласно которой из $|f(x)| \leqslant \varphi(x)$ при

---

[1]) Элементарно это доказывается так: пусть $0 < x < 1$, $\frac{1}{x} = z$, $z - 1 = y$ (> 0); тогда $z^n = (1 + y)^n > 1 + ny + \frac{n(n-1)}{2}y^2 > \frac{n(n-1)}{2}y^2$, и, значит, $nx^n = \frac{n}{z^n} < \frac{2}{(n-1)y^2} \to 0$ при $n \to \infty$.

$a \leqslant x \leqslant b$ вытекает

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b \varphi(x)\,dx,$$

мы находим поэтому

$$\left|\int_a^b r_n(x)\,dx\right| \leqslant \varepsilon(b-a) \qquad (n \geqslant n_0);$$

ввиду произвольности ε это и показывает, что

$$\int_a^b r_n(x)\,dx \to 0 \qquad (n \to \infty),$$

а это равносильно равенству (13).

Равномерная сходимость, будучи, таким образом, достаточной предпосылкой почленной интегрируемости ряда непрерывных функций, не является, как и в предыдущей проблеме, необходимым условием этой интегрируемости. Чтобы в этом убедиться, достаточно снова рассмотреть пример, приведенный на стр. 115; здесь сходимость неравномерна, но в то же время при $n \to \infty$

$$\int_0^1 r_n(x)\,dx = \int_0^1 x^{2n}dx - \int_0^1 x^n dx = \frac{1}{2n+1} - \frac{1}{n+1} \to 0,$$

т. е. имеет место равенство (13).

Равномерная сходимость конкретно заданных рядов чаще всего устанавливается с помощью следующего простого признака: *если* $\sum_{n=1}^{\infty} \alpha_n$ — *сходящийся ряд с положительными членами и если существует такое* n_0, *что для всех* $n \geqslant n_0$ *и всех* $x \in M$

$$|u_n(x)| \leqslant \alpha_n,$$

то ряд (12) *равномерно сходится на множестве* M. В самом деле, в условиях теоремы при достаточно

большом n и любом $k > 0$ для всех $x \in M$

$$|u_{n+1}(x) + u_{n+2}(x) + \ldots + u_{n+k}(x)| \leqslant \leqslant \sum_{i=n+1}^{n+k} |u_i(x)| \leqslant \sum_{i=n+1}^{n+k} \alpha_i < \varepsilon$$

(последнее неравенство выполняется в силу критерия Коши для всех достаточно больших n при любом $k > 0$), откуда в силу критерия Коши ряд (12) сходится и при достаточно большом n для всех $x \in M$

$$|r_n(x)| \leqslant \varepsilon,$$

что и требовалось установить.

Степенные ряды. Бесспорно, наиболее важный для анализа специальный класс функциональных рядов составляют так называемые *степенные ряды*, т. е. ряды вида

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots, \qquad (14)$$

где $a_0, a_1, a_2, \ldots, a_n, \ldots$ суть данные вещественные числа. Мы рассмотрим здесь некоторые важнейшие свойства этого класса рядов.

Вы знаете, прежде всего, что для степенного ряда всегда существует некоторый открытый отрезок $(-r, r)$ с центром в точке 0, так называемый *интервал сходимости*, внутри которого данный ряд сходится, а вне — расходится; число r называется *радиусом сходимости* данного ряда; r может иметь любое значение от 0 до $+\infty$, включая оба этих крайних значения (так, $r = 0$ при $a_n = n!$, $r = +\infty$ при $a_n = \frac{1}{n!}$); заданием радиуса сходимости область сходимости данного ряда определяется с точностью до двух граничных ее точек, именно, область сходимости может быть либо закрытым, либо открытым отрезком, либо, наконец, «полузакрытым», т. е. содержащим один из своих концов и не содержащим другого; так, каждый из трех рядов с коэффициентами соответственно $a_n = 1$, $a_n = \frac{1}{n+1}$, $a_n = \frac{1}{(n+1)^2}$ имеет радиус сходимости, равный единице. Первый из этих рядов расходится при $x = 1$ и $x = -1$, т. е. в

обоих концах интервала сходимости (этот ряд есть геометрическая прогрессия со знаменателем x), второй ряд при $x = 1$ обращается в гармонический ряд, а при $x = -1$ в «ряд Лейбница»

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} + \ldots;$$

он, следовательно, расходится при $x = 1$ и условно сходится при $x = -1$; третий ряд абсолютно сходится в обоих концах интервала сходимости.

Напомним вкратце, как устанавливается эта специальная форма области сходимости для степенных рядов. Вывод ее основан на замечательной теореме: *если ряд* (14) *сходится при* $x = \alpha$, *то он абсолютно сходится при любом значении* x, *для которого* $|x| < \alpha$. Вот ее доказательство: из сходимости ряда (14) при $x = \alpha$ вытекает, что $a_n\alpha^n \to 0$ при $n \to \infty$; значит, существует такое $c > 0$, что $|a_n\alpha^n| < c$ $(n=1, 2, \ldots)$; пусть теперь $|x| < |\alpha|$, тогда

$$\sum_{k=1}^{n} |a_k x^k| = \sum_{k=1}^{n} |a_k a^k| \cdot \left|\frac{x}{\alpha}\right|^k < c \sum_{k=1}^{n} \left|\frac{x}{\alpha}\right|^k <$$

$$< c \sum_{k=1}^{\infty} \left|\frac{x}{\alpha}\right|^k = \frac{c}{1 - \left|\frac{x}{\alpha}\right|};$$

так как правая часть от n не зависит, то левая остается ограниченной при $n \to \infty$, т. е. ряд (14) абсолютно сходится. Из этой теоремы сразу вытекает специальная форма области сходимости для степенных рядов, так как вместе с точкой x этой области принадлежат и все точки, лежащие к нулю ближе, чем x. Другим следствием этой теоремы является то, что в каждой *внутренней* точке интервала сходимости ряд (14) сходится *абсолютно*. Что касается равномерной сходимости, то для всего интервала сходимости мы ее утверждать не можем, как показывает уже пример геометрической прогрессии ($a_n = 1$): как бы велико ни было n, $r_n(x) = \sum_{k=n+1}^{\infty} x^k$ становится сколь угодно большим, если x

достаточно близко к единице. Но легко показать, что *степенной ряд сходится равномерно на любом отрезке, концы которого лежат внутри интервала сходимости.* В самом деле, пусть r есть радиус сходимости ряда (14) и пусть $0 < r' < r$; для всех x, удовлетворяющих неравенству $|x| \leqslant r'$, мы имеем при любом n

$$|a_n x^n| \leqslant |a_n| r'^n;$$

а так как $\sum_{n=1}^{\infty} |\alpha_n| r'^n$ есть сходящийся ряд с положительными членами, то ряд (14) равномерно сходится на отрезке $|x| \leqslant r'$ (см. признак на стр. 121).

Из этой теоремы, между прочим, вытекает то весьма важное следствие, что *сумма степенного ряда есть функция непрерывная в каждой внутренней точке интервала сходимости.*

В случае, когда ряд (14) сходится при $x = r$, можно, впрочем, утверждать, что он равномерно сходится на отрезке $[0, r]$ (и, значит, сумма его непрерывна на этом отрезке); эта замечательная *теорема Абеля* играет важную роль в теории функций.

Для доказательства установим сначала следующую элементарную лемму, также принадлежащую Абелю:

$$\sum_{n=n_1}^{n_2} a_n (b_n - b_{n-1}) = \sum_{n=n_1}^{n_2} b_n (a_n - a_{n+1}) - \\ - a_{n_1} b_{n_1-1} + a_{n_2+1} b_{n_2},$$

где $n_2 > n_1 > 0$, а a_k и b_k — любые числа.

В самом деле,

$$\sum_{n=n_1}^{n_2} a_n (b_n - b_{n-1}) =$$
$$= a_{n_1} (b_{n_1} - b_{n_1-1}) + a_{n_1+1} (b_{n_1+1} - b_{n_1}) +$$
$$+ a_{n_1+2} (b_{n_1+2} - b_{n_1+1}) + \ldots + a_{n_2-1} (b_{n_2-1} - b_{n_2-2}) +$$
$$+ a_{n_2} (b_{n_2} - b_{n_2-1}) = - a_{n_1} b_{n_1-1} + b_{n_1} (a_{n_1} - a_{n_1+1}) +$$
$$+ b_{n_1+1} (a_{n_1+1} - a_{n_1+2}) + \ldots + b_{n_2-1} (a_{n_2-1} - a_{n_2}) +$$
$$+ b_{n_2} a_{n_2} = \sum_{n=n_1}^{n_2} b_n (a_n - a_{n+1}) - a_{n_1} b_{n_1-1} + a_{n_2+1} b_{n_2}.$$

Теперь переходим к доказательству самой теоремы.

Теорема Абеля. *Если ряд* (14) *сходится при* $x = r > 0$, *то он равномерно сходится на отрезке* $0 \leqslant x \leqslant r$.

Доказательство. Пусть ε — произвольное положительное число; в силу критерия Коши при всех достаточно больших n и любом $k > 0$

$$\left| \sum_{i=n+1}^{n+k} a_i r^i \right| < \varepsilon;$$

положим для краткости $\sum\limits_{i=n+1}^{n+k} a_i r^i = \sigma_k$ $(\sigma_0 = 0)$, так что $|\sigma_k| < \varepsilon$ $(k = 0, 1, 2, \ldots)$. Тогда при $0 \leqslant x \leqslant r$

$$\sum_{i=n+1}^{n+k} a_i x^i = \sum_{i=n+1}^{n+k} a_i r^i \left(\frac{x}{r}\right)^i = \sum_{i=1}^{k} a_{i+n} r^{i+n} \left(\frac{x}{r}\right)^{i+n} =$$

$$= \sum_{i=1}^{k} (\sigma_i - \sigma_{i-1}) \left(\frac{x}{r}\right)^{i+n},$$

а потому в силу доказанной леммы (и ввиду равенства $\sigma_0 = 0$)

$$\left| \sum_{i=n+1}^{n+k} a_i x^i \right| =$$

$$= \left| \sum_{i=1}^{k} \sigma_i \left\{ \left(\frac{x}{r}\right)^{n+i} - \left(\frac{x}{r}\right)^{n+i+1} \right\} + \sigma_k \left(\frac{x}{r}\right)^{n+k+1} \right| \leqslant$$

$$\leqslant \sum_{i=1}^{k} |\sigma_i| \left(\frac{x}{r}\right)^{n+i} \left(1 - \frac{x}{r}\right) + |\sigma_k| \left(\frac{x}{r}\right)^{n+k+1} \leqslant$$

$$\leqslant \varepsilon \left\{ \left(1 - \frac{x}{r}\right) \sum_{i=1}^{k} \left(\frac{x}{r}\right)^{n+i} + \left(\frac{x}{r}\right)^{n+k+1} \right\} = \varepsilon \left(\frac{x}{r}\right)^{n+1} \leqslant \varepsilon;$$

в силу критерия Коши мы заключаем отсюда, что ряд (14) сходится; переходя же к пределу при $k \to \infty$, получаем из последнего неравенства, что при достаточно

большом n

$$|r_n(x)| \leqslant \varepsilon \qquad (0 \leqslant x \leqslant r);$$

это и означает, что ряд (14) равномерно сходится на отрезке $[0, r]$.

Необходимо отметить, однако, что из сходимости ряда (14) при $x = r$ неправильно было бы заключать о его равномерной сходимости на всем отрезке $(-r, r)$; так, ряд

$$1 - \frac{x}{1} + \frac{x^2}{2} - \frac{x^3}{3} + \ldots,$$

сходящийся при $x = 1$, очевидно, расходится при $x = -1$ и не может сходиться равномерно на всем интервале сходимости $(-1, 1)$.

Непосредственно ясно, что одна из важнейших задач, связанных со степенными рядами, состоит в определении радиуса сходимости r по данным коэффициентам ряда a_n; как вы, вероятно, знаете, элементарные признаки сходимости числовых рядов с положительными членами дают возможность решить эту задачу в некоторых частных предположениях; так, например, если существует $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = l$ при $n \to \infty$, то $r = \frac{1}{l}$, $+\infty$ или 0, если, соответственно, $0 < l < +\infty$, $l = 0$ или $l = +\infty$. Можно, однако, указать метод, позволяющий решить эту задачу во всех случаях.

Теорема. *Положим* $\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|} = l$. *Тогда*

$$r = \begin{cases} \frac{1}{l}, & \text{если} \quad 0 < l < +\infty, \\ 0, & \text{если} \quad l = +\infty, \\ +\infty, & \text{если} \quad l = 0. \end{cases}$$

Так как верхний предел (конечный или бесконечный) существует у всякой последовательности, то этой теоремой поставленная задача действительно решается во всех случаях.

Прежде чем перейти к доказательству, вспомним: если число l есть верхний предел некоторой последовательности, то при любом $\varepsilon > 0$ с некоторого места все члены этой последовательности меньше чем $l+\varepsilon$; напротив, как угодно далеко в данной последовательности будут встречаться числа, бо́льшие чем $l-\varepsilon$.

Доказательство теоремы. 1) Пусть $0<l<$ $<+\infty$; положим $r=\frac{1}{l}$; нужно доказать, что ряд (14) сходится при $0<x<r$ и расходится при $x>r$.

1a) Пусть $0<x<r=\frac{1}{l}$, так что $lx<1$; очевидно, можно выбрать число $\varepsilon>0$ столь малым, что $(l+\varepsilon)x<1$; из определения числа l вытекает, что при достаточно большом n

$$\sqrt[n]{|a_n|}<l+\varepsilon, \quad |a_n|<(l+\varepsilon)^n$$

и, следовательно,

$$|a_n|x^n<[(l+\varepsilon)x]^n;$$

так как $(l+\varepsilon)x<1$, то члены ряда (14) по абсолютной величине при достаточно большом n не превосходят соответственных членов сходящейся геометрической прогрессии, откуда следует сходимость ряда (14).

1b) Пусть $x>r=\frac{1}{l}$, так что $lx>1$; при достаточно малом $\varepsilon>0$ мы будем иметь $(l-\varepsilon)x>1$. По определению числа l существуют сколь угодно большие значения n, для которых $\sqrt[n]{|a_n|}>l-\varepsilon$, откуда $|a_n|>$ $>(l-\varepsilon)^n$ и, следовательно,

$$|a_n|x^n>[(l-\varepsilon)x]^n>1.$$

Отсюда следует, что ряд (14) содержит бесконечное множество членов, абсолютная величина которых превосходит единицу; общий член его не может быть поэтому бесконечно малым, и ряд расходится.

2) Пусть $l=0$; требуется доказать, что ряд (14) сходится для любого $x>0$. По определению числа l

при достаточно большом n

$$\sqrt[n]{|a_n|} < \frac{1}{2x}, \quad |a_n| < \frac{1}{2^n x^n}, \quad |a_n| x^n < \frac{1}{2^n},$$

т. е. n-й член ряда (14) при достаточно большом n по абсолютному значению снова меньше n-го члена некоторой сходящейся прогрессии; значит, ряд (14) сходится.

3) Пусть $l = +\infty$; требуется доказать, что ряд (14) расходится при любом $x > 0$. По определению l мы для сколь угодно больших n должны иметь

$$\sqrt[n]{|a_n|} > \frac{1}{x}, \quad |a_n| x^n > 1;$$

снова, значит, общий член ряда (14) не может быть бесконечно малым, и ряд расходится. Таким образом, теорема доказана во всех случаях.

Л Е К Ц И Я V

ПРОИЗВОДНАЯ

Производная функция и производные числа.— Дифференциал.— Теорема Лагранжа.— Повторное дифференцирование.— Пределы отношений бесконечно малых и бесконечно больших.— Формула Тейлора.— Экстремальные значения.— Частное дифференцирование.— Неявные функции.

Производная функция и производные числа. До сих пор мы преимущественно имели дело с возведением вспомогательных аналитических конструкций или с анализом основоположных понятий. Сегодня мы переходим в центральное здание математического анализа, которое составляют учения о дифференцировании и интегрировании.

Пусть $y = f(x)$ есть функция переменной x, определенная в некоторой окрестности точки $x = \alpha$. Если мы от этой точки перейдем к точке $\alpha + h$, то функция y получит «приращение» $f(\alpha + h) - f(\alpha)$. Если мы хотим получить представление о том, как быстро меняется величина y при изменении независимой переменной x, насколько «чувствительна» функция $f(x)$ к такому изменению, то мы должны, конечно, так или иначе сопоставить, сравнить между собой приращение функции y и то изменение h величины x, которое повлекло за собой это приращение; естественнее всего рассматривать с этой целью отношение

$$\frac{f(\alpha + h) - f(\alpha)}{h}, \tag{1}$$

которое дает нам среднее приращение функции y, рассчитанное на единицу приращения величины x; однако расчет этот произведен при определенном значении h

и будет, вообще говоря, давать различные результаты при различных h; для того чтобы поставленная задача получила однозначное решение, мы должны будем условиться выбирать величину h на основе какого-либо единого принципа.

Если нашей целью является исследование поведения функции $f(x)$ в ближайшем соседстве точки α, то ясно, что величина (1) в качестве меры «изменчивости» функции y будет тем больше удовлетворять нашим запросам, чем меньше мы выберем $|h|$; в самом деле, ведь величина (1) характеризует нам «среднюю изменчивость» функции на отрезке $[\alpha, \alpha + h]$, который тем ближе примыкает к точке α, чем меньше $|h|$. От этих соображений для нас, знакомых с понятием предельного перехода, уже всего один шаг до признания, что мы получим наиболее удовлетворительное решение нашей задачи, если за характеристику изменчивости нашей функции в ближайшем соседстве точки α примем *предел величины* (1) *при* $h \to 0$ (в предположении, разумеется, что этот предел существует), т. е. величину

$$f'(\alpha) = \lim_{h \to 0} \frac{f(\alpha + h) - f(\alpha)}{h};$$

другая общепринятая система обозначений есть, как вы знаете,

$$h = \Delta x, \quad f(\alpha + h) - f(\alpha) = \Delta y, \quad f'(\alpha) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Эту последнюю величину мы называем *производной* функции $f(x)$ в точке α (или «при $x = \alpha$»). Таким образом, производная данной функции в данной точке призвана характеризовать собой относительную изменчивость этой функции в непосредственном соседстве данной точки; чем больше $|f'(\alpha)|$, тем чувствительнее величина y к малым отклонениям величины x от ее начального значения α; знак же величины $f'(\alpha)$ характеризует направление этой изменчивости: величина $f'(\alpha)$ положительна или отрицательна в зависимости от того, возрастает или убывает функция y при малых увеличениях, которые претерпевает величина x, отправляясь

от начального значения α. Если функцию $y = f(x)$ изобразить графически, то интересующая нас изменчивость иллюстрируется, очевидно, тем, насколько круто поднимается или опускается начерченная кривая, когда переменная величина x проходит через значение α; в точных терминах, как вы знаете, производная иллюстрируется угловым коэффициентом касательной к кривой $y = f(x)$ при $x = \alpha$.

Наиболее простое реальное истолкование производная получает в том случае, когда независимая переменная x символизирует *время*. При этом условии величина (1) представляет собой среднюю скорость изменения величины y в течение промежутка времени $[\alpha, \alpha + h]$, а производная $f'(\alpha)$ — «истинную скорость» этого изменения в момент α; в частности, если $y = f(x)$ означает путь, пройденный движущейся точкой за время от некоторого постоянного момента a до момента x, понятие производной в точности совпадает с механическим понятием мгновенной скорости.

В математическом естествознании и других приложениях анализа производная играет такую выдающуюся роль именно потому, что она дает локальную характеристику изучаемого явления в наиболее важном отношении — количественно оценивает изменяемость одной из двух связанных между собой переменных величин при изменении другой.

Если бы вы сегодня здесь слышали о производной в первый раз, вы, вероятно, теперь спросили бы: почему же говорят «производная», откуда этот женский род? Ведь речь идет о пределе, числе и т. д. Но вы, конечно, знаете, в чем дело: «производная» сокращенно означает «производная функция»; весь расчет и все рассуждения, проведенные нами для одной (произвольно выбранной) точки α отрезка $[a, b]$, мы можем проделать для любой точки x этого отрезка (в предположении, разумеется, что каждый раз вычисляемый предел существует); получаемая таким образом функция $f'(x)$ и называется *производной* функцией. Само собой разумеется, все это рассуждение ничего не меняет в том основном факте, что производная есть *локальная* характеристика данной функции, особо вычисляемая для

каждой отдельной точки и призванная описывать собой поведение этой функции не на целом отрезке $[a, b]$, а лишь в ближайшем соседстве его отдельных точек.

Функция $f(x)$ называется *дифференцируемой* при $x = \alpha$, если она имеет производную в точке α; она *дифференцируема внутри отрезка* $[a, b]$, если производная существует в каждой внутренней точке этого отрезка. Очевидно, что дифференцируемость есть, подобно непрерывности, локальное свойство функций. Вы знаете, конечно, что дифференцируемой в точке α может быть только функция, непрерывная в этой точке; это видно из того, что когда знаменатель дроби (1) стремится к нулю, то для существования предела необходимо, чтобы стремился к нулю и числитель этой дроби, а это как раз и выражает непрерывность функции $f(x)$ в точке α. Вы знаете, вероятно, также, что обратное заключение неверно: непрерывная функция может и не иметь производной; так, функция $f(x)$, которая изображена на рис. 19 и для которой нет надобности приводить аналитическое выражение, очевидно, непрерывна; в точке α существует предел выражения (1) при $h \to +0$ и при $h \to -0$, но эти два предела различны между собой (геометрически это сказывается в том, что в точке α кривая не имеет единой касательной; можно надеяться, что после нашей третьей лекции никто из вас не станет утверждать, что рис. 19 изображает «не одну, а две функции»). Поэтому предел при $h \to 0$ не существует. Пределы выражения (1) при $h \to +0$ и $h \to -0$, если они существуют, называются соответственно правой и левой производной функции $f(x)$ в точке α. Для дифференцируемости функции $f(x)$ в точке α, очевидно, необходимо и достаточно, чтобы правая и левая производные в этой точке существовали и совпадали между собой.

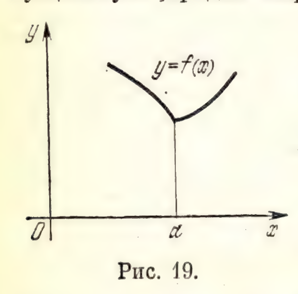


Рис. 19.

Рис. 19 дает нам простейший пример случая, когда непрерывная функция в отдельной точке лишена производной; однако здесь все же существуют правая и левая

производные; естественно поставить вопрос, всегда ли дело обстоит так; и легко убедиться, что бывают случаи, когда непрерывная функция недифференцируема в гораздо более глубоком смысле. Рис. 20 схематически изображает поведение функции $y = x \sin\frac{1}{x}$ в соседстве точки $x = 0$; так как при $x \to 0$

$$\left| x \sin \frac{1}{x} \right| < |x| \to 0,$$

то, полагая дополнительно $y = 0$ при $x = 0$, мы видим,

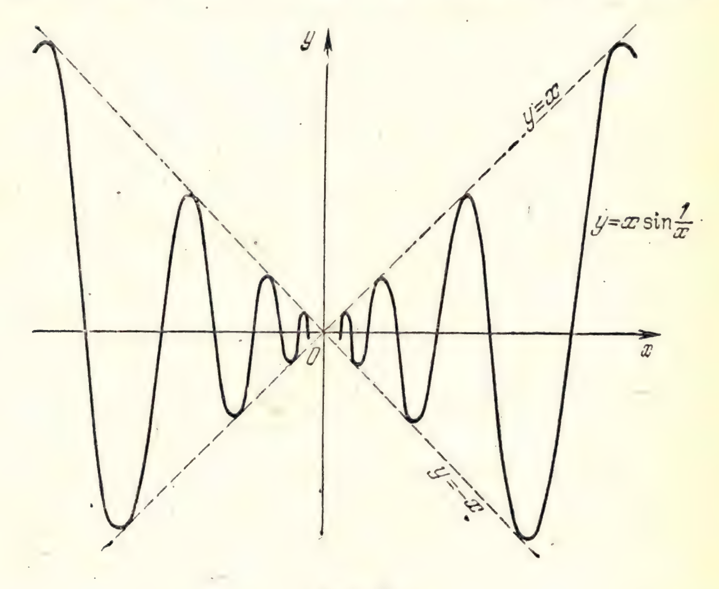


Рис. 20.

что функция y непрерывна в точке 0; но так как при стремлении x к нулю (например, справа) $\frac{1}{x} \to +\infty$ и потому $\sin\frac{1}{x}$ бесчисленное множество раз колеблется между 1 и -1, то $x \sin \frac{1}{x}$ совершает бесчисленное множество колебаний между прямыми $y = x$ и $y = -x$;

выражение (1), равное в нашем случае $\sin\frac{1}{x}$ (при $\alpha=0$ и $h=x$), имеет частичными пределами все числа отрезка $[-1, 1]$; верхний предел его есть 1, нижний -1; в точности то же самое мы имеем при $x\to-0$; таким образом, здесь в точке 0 нет ни правой, ни левой производной.

В общем случае выражение (1) имеет верхний и нижний пределы как при $h\to+0$, так и при $h\to-0$; эти четыре предела называют *производными числами* данной функции в данной точке; каждое из них может быть либо числом, либо одним из символов $\pm\infty$. Таким образом, любая функция в любой точке (в некоторой окрестности которой она определена) имеет четыре производных числа: правое верхнее, правое нижнее, левое верхнее, левое нижнее; если оба правых (левых) производных числа совпадают, функция имеет в данной точке правую (левую) производную; в случае, когда все четыре числа конечны и совпадают между собой (и только в этом случае), функция дифференцируема в данной точке. Пример случая, когда все четыре производных числа в некоторой точке бесконечны, дает функция

$$y=\begin{cases}\sqrt{|x|}\sin\frac{1}{x} & \text{при } x\neq 0,\\ 0 & \text{при } x=0\end{cases}$$

в точке $x=0$, как вы теперь легко разберетесь сами.

Если теперь принять во внимание, что только что нарисованная нами картина может для одной и той же функции в разных точках складываться совершенно поразному, то вы поймете, какое сложное образование может являть собой одна и та же функция в отношении к идее дифференцирования. Явления, подобные изображенному на рис. 20, отнюдь не всегда являются принадлежностью лишь отдельных, изолированных точек; существуют непрерывные функции, имеющие столь же сложную структуру в соседстве *каждой* точки (и, в частности, значит, нигде не дифференцируемые); к сожалению, рамки наших лекций не позволяют нам остановиться на каком-либо из построенных теперь уже в большом числе примеров подобных функций.

Вместо этого рассмотрим пример применения производных чисел. Вы знаете, конечно, что для дифференцируемых функций поведение их в смысле возрастания и убывания тесно связано со знаком производной функции. В частности, если $f'(x) \geqslant 0$ ($\leqslant 0$) во всех точках отрезка $[a, b]$, то функция $f(x)$ — неубывающая (невозрастающая) на этом отрезке. Эти признаки возрастания и убывания не оставляют, конечно, желать ничего лучшего в тех случаях, когда они применимы, т. е. когда данная функция имеет производную; но в случае недифференцируемых функций признаки этого рода, разумеется, не дают ничего; а между тем при ближайшем исследовании оказывается, что идея подобного рода признаков возрастания и убывания функций ни в какой степени не стоит в зависимости от их дифференцируемости; чтобы в этом убедиться, докажем следующее предложение.

Теорема. *Пусть $f(x)$ — непрерывная функция на отрезке $[a, b]$, и пусть одно из четырех производных чисел этой функции — обозначим его через $Df(x)$ — неотрицательно для всех $a \leqslant x \leqslant b$; тогда $f(b) \geqslant f(a)$.*

Само собой понятно, что эта редакция признака существенно сильнее обычной, так как здесь речь идет о любой непрерывной функции, вообще говоря, недифференцируемой.

Доказательство. Пусть, в противоречии с утверждением теоремы, $f(b) < f(a)$ и число ε таково, что $\frac{f(a)-f(b)}{b-a} > \varepsilon > 0$; пусть для определенности $Df(x)$ — правое верхнее производное число функции $f(x)$. Положим $\varphi(x) = f(x) - f(a) + \varepsilon(x-a)$; очевидно, $\varphi(a) = 0$ и

$$\varphi(b) = (b-a)\left(\varepsilon - \frac{f(a)-f(b)}{b-a}\right) < 0. \qquad (2)$$

Пусть M — множество точек отрезка $[a, b]$, в которых $\varphi(x) = 0$, и α — верхняя грань этого множества; если бы мы имели $\varphi(\alpha) > 0$ (или $\varphi(\alpha) < 0$), то точку α можно было бы окружить окрестностью U, в любой точке x которой $\varphi(x) > 0$ (или соответственно $\varphi(x) < 0$); с другой стороны, в силу определения верхней грани

любая окрестность точки α должна содержать точки, где $\varphi(x) = 0$. Это противоречие показывает, что $\varphi(\alpha) = 0$. Так как $\varphi(b) < 0$, то, очевидно, $\varphi(x) < 0$ для $\alpha < x \leqslant b$; для любого такого x

$$\frac{\varphi(x) - \varphi(\alpha)}{x - \alpha} < 0,$$

вследствие чего и $D\varphi(\alpha) \leqslant 0$. Но

$$D\varphi(\alpha) = Df(\alpha) + \varepsilon,$$

откуда

$$Df(x) = D\varphi(\alpha) - \varepsilon < 0,$$

что стоит в противоречии с условием теоремы, которая, таким образом, доказана.

Дифференциал. Следующим основным понятием учения о дифференцировании служит понятие *дифференциала.* В настоящее время мы считаем дифференциал вторичным понятием, определяемым через понятие производной; однако это не всегда было так; в эпоху зарождения анализа бесконечно малых и еще долгое время спустя первичным понятием анализа считали именно дифференциал, производную же определяли как отношение дифференциалов, т. е. как вторичное понятие. При этом часто понятие дифференциала оставалось без четкого определения и даже несло в себе противоречивые черты, так как в эту эпоху в математике диалектические приемы мышления, способные охватить в качестве единого объекта переменную величину в процессе ее изменения, не получили еще достаточного развития.

Формальное определение дифференциала вам, конечно, известно: мы называем дифференциалом функции $y = f(x)$ величину

$$dy = f'(x)\Delta x, \tag{3}$$

где Δx — приращение независимой переменной; дифференциал функции $f(x)$ есть, таким образом, функция двух независимых между собой переменных — величины x и ее приращения Δx; значения этих двух переменных никак между собой не связаны и могут быть выбраны совершенно независимо друг от друга.

Рассматривая, в частности, функцию $y = x$, мы приходим к выводу, что $dx = \Delta x$, т. е. для независимой переменной дифференциал и приращение всегда совпадают между собой; подставляя в формулу (3) dx на место Δx, мы находим

$$dy = f'(x)\,dx,$$

откуда

$$y' = f'(x) = \frac{dy}{dx},$$

т. е. производная равна отношению дифференциала функции к дифференциалу независимой переменной.

Эти чисто формальные рассмотрения не позволяют еще, однако, понять то огромное значение, какое имеет понятие дифференциала для анализа и его приложений. Чтобы разобраться в этом вопросе, надо вглядеться в идею дифференциала по существу. Будет, может быть, всего убедительнее начать с рассмотрения того частного случая, когда x означает время, а $y = f(x)$ — путь, пройденный движущейся точкой за промежуток времени от 0 до x. В этом случае, как мы знаем, производная $f'(x)$ означает мгновенную скорость движущейся точки в момент x; следовательно, $dy = f'(x)\,\Delta x$ есть величина того пути, который прошла бы движущаяся точка за промежуток времени $[x, x + \Delta x]$, если бы в течение этого промежутка она двигалась равномерно с той скоростью, какую она имеет в начале этого промежутка. Действительное расстояние, пройденное точкой за этот промежуток времени, будет, вообще говоря, другим, так как скорость не будет оставаться постоянной.

В общем случае, как мы знаем, производная $f'(x)$ может рассматриваться как мера относительной изменчивости величин y и x при данном значении x; дифференциал $dy = f'(x)\,\Delta x$ мы можем поэтому рассматривать как то приращение, которое получила бы функция y (при изменении величины x от ее первоначального значения до $x + \Delta x$), если бы во всех точках отрезка $[x, x + \Delta x]$ эта мера оставалась такой же, какова она в начале этого отрезка. Эту концепцию наглядно иллюстрирует рис. 21: dy есть приращение,

которое получила бы на пути от x до $x + \Delta x$ ордината кривой $y = f(x)$, если бы наклон кривой на всем этом промежутке был такой же, как в точке x, т. е. если бы мы эту кривую заменили касательной, проведенной к ней в точке, абсцисса которая равна x.

Вероятно, многим из вас известно, что в прикладных науках при малых значениях Δx часто не делают различия между приращением Δy и дифференциалом dy

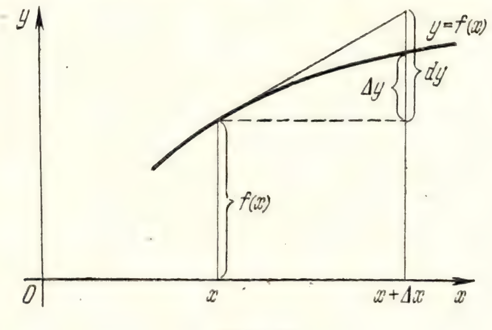


Рис. 21.

функции $y = f(x)$; иногда это дает даже повод к неправильному и вредному представлению о дифференциале как о «бесконечно малом приращении» (в действительности дифференциал, вообще говоря, не является приращением). Нам полезно будет постараться ответить на два вопроса, естественно встающих по поводу такой замены: 1) в какой мере эта замена обоснована? и 2) какую пользу может она принести?

Чтобы ответить на первый вопрос, будем исходить из соотношения

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x};$$

обозначим через α бесконечно малую в силу этого соотношения разность $\frac{\Delta y}{\Delta x} - f'(x)$; это дает

$$\Delta y = f'(x)\,\Delta x + \alpha\,\Delta x = dy + \alpha\,\Delta x. \qquad (3')$$

Так как $\alpha \to 0$ при $\Delta x \to 0$, то произведение $\alpha\,\Delta x$ есть

бесконечно малая высшего порядка сравнительно с Δx; но отношение $\frac{dy}{dx} = f'(x)$ есть величина постоянная, а отношение $\frac{\Delta y}{\Delta x}$ при $\Delta x \to 0$ имеет эту постоянную величину своим пределом; поэтому, если только $f'(x) \neq 0$, все три величины Δx, Δy, dy имеют одинаковый порядок малости, так что величина $\alpha \Delta x$, будучи бесконечно малой высшего порядка сравнительно с Δx, есть вместе с тем бесконечно малая высшего порядка и относительно каждой из величин Δy, dy; таким образом, полученное нами соотношение показывает, что, если $f'(x) \neq 0$, *разность между приращением и дифференциалом функции при $\Delta x \to 0$ есть бесконечно малая высшего порядка по сравнению с каждой из этих величин в отдельности*; другими словами, заменяя приращение дифференциалом (или обратно), мы допускаем лишь бесконечно малую *относительную* погрешность.

Полученное нами соотношение при $f'(x) \neq 0$, очевидно, равносильно соотношению

$$\frac{\Delta y}{dy} = 1 + \frac{\alpha}{f'(x)} \to 1 \qquad (\Delta x \to 0),$$

показывающему, что Δy и dy — эквивалентные бесконечно малые при $\Delta x \to 0$. На этих результатах и основывается возможность замены приращения функции при малых Δx, в порядке приближения, ее дифференциалом.

Чтобы ответить на второй поставленный вопрос — какую пользу может принести замена приращения функции ее дифференциалом? — заметим, что строение дифференциала теоретически проще и практически удобнее, чем строение приращения; дифференциал dy есть *линейная* функция величины Δx, характер ее изменения при изменении Δx исключительно прост, и изучение его не требует ничего, кроме вычисления значения $f'(x)$ *в одной точке*; разумеется, ничего похожего мы не имеем для величины Δy. Представьте себе, например, что вы хотите составить таблицу значений функции $y = \sin x$ при значениях x, очень близких к 60°, например: 60°01′, 60°02′ и т. д. У вас нет средств

для отыскания точной величины этих значений. Вы знаете, что при $x = 60°$ $y = \frac{\sqrt{3}}{2}$, но при переходе к другим, близлежащим значениям x у вас нет метода для отыскания соответствующих приращений Δy величины y; теперь посмотрите, что получится, если вы, пользуясь малостью приращений Δx в вашей задаче, решитесь заменить приращения Δy дифференциалами dy; так как $y' = \cos 60° = \frac{1}{2}$, то $dy = \frac{1}{2}\Delta x$, где Δx выражено в радианной мере $\left(\operatorname{arc} 1' = \frac{\pi}{180 \cdot 60}\right)$, вследствие чего мы сразу получаем

$$\sin 60°01' \approx \frac{\sqrt{3}}{2} + \frac{1}{2} \cdot \frac{\pi}{180 \cdot 60},$$

$$\sin 60°02' \approx \frac{\sqrt{3}}{2} + \frac{1}{2} \cdot \frac{2\pi}{180 \cdot 60},$$

$$\sin 60°03' \approx \frac{\sqrt{3}}{2} + \frac{1}{2} \cdot \frac{3\pi}{180 \cdot 60},$$

. . . .

Как мы видели, дифференциал dy обладает следующими двумя замечательными свойствами: 1) он есть линейная функция от Δx, и 2) он отличается от Δy на бесконечно малую высшего порядка по сравнению с Δx. Покажем теперь, что этими двумя свойствами дифференциал полностью определяется, так что можно было бы начинать учение о дифференциале именно с такого определения (хотя, конечно, для доказательства *существования* дифференциала у всякой дифференцируемой функции все равно нельзя было бы избежать введения производной).

В самом деле, пусть $dy = a\,\Delta x$, где a не зависит от Δx, и пусть $\Delta y - dy = \alpha\,\Delta x$, где $\alpha \to 0$ при $\Delta x \to 0$. Тогда $\Delta y = dy + \alpha\,\Delta x = (a + \alpha)\,\Delta x$, откуда

$$\frac{\Delta y}{\Delta x} = a + \alpha, \quad \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x) = a, \quad dy = f'(x)\,\Delta x,$$

что и требовалось доказать.

Из других свойств дифференциала важное значение имеет (но во многих руководствах не оттеняется в достаточной мере) так называемая *инвариантность* его относительно любой замены независимой переменной. Вот что это значит. Если x есть независимая переменная, то, как мы видели,

$$dy = f'(x)\,\Delta x = f'(x)\,dx; \tag{4}$$

пусть теперь мы устанавливаем взгляд на величину x как на функцию $x = \varphi(t)$ новой переменной t; ясно, что при этом оба соотношения $dy = f'(x)\Delta x$ и $dy = f'(x)\,dx$ не могут остаться справедливыми, так как теперь уже (вообще говоря) $dx = \varphi'(t)\Delta t \neq \Delta x$.

Замечательно, что второе соотношение, т. е.

$$dy = f'(x)\,dx,$$

остается в силе при любой такой замене; в самом деле, в результате этой замены y становится функцией от t:

$$y = \psi(t) = f[\varphi(t)];$$

отсюда по правилу дифференцирования функции от функции

$$\psi'(t) = f'[\varphi(t)]\varphi'(t),$$
$$dy = \psi'(t)\,dt = f'[\varphi(t)]\varphi'(t)\,dt,$$

а так как $\varphi(t) = x$, $\varphi'(t)\,dt = dx$, то, действительно, $dy = f'(x)\,dx$, как если бы x была независимая переменная. Таким образом, в выражении производной $y' = \frac{dy}{dx}$ безразлично, есть ли x независимая переменная или дифференцируемая функция какой-нибудь другой переменной.

Мы только что пользовались правилом дифференцирования функции от функции. Вообще говоря, мы в этих лекциях не собираемся говорить об элементарных правилах дифференцирования; но именно этому правилу стоит уделить немного внимания, потому что в большинстве руководств оно выводится либо неправильно, либо излишне сложно; вот его простой вывод: пусть $y = f(x)$, $x = \varphi(t)$, $y = \psi(t) = f[\varphi(t)]$; функции f

и φ предполагаются, разумеется, дифференцируемыми, так что в соотношении (3′) $\alpha \to 0$ при $\Delta x \to 0$; из этого соотношения

$$\frac{\Delta y}{\Delta t} = f'(x)\frac{\Delta x}{\Delta t} + \alpha\frac{\Delta x}{\Delta t};$$

но при $\Delta t \to 0$ мы имеем $\alpha \to 0$, $\frac{\Delta x}{\Delta t} \to \varphi'(t)$ и $\frac{\Delta y}{\Delta t} \to \psi'(t)$, вследствие чего в пределе

$$\psi'(t) = f'(x)\varphi'(t) = f'[\varphi(t)]\varphi'(t);$$

это и есть правило, которое мы имели в виду вывести.

Теорема Лагранжа. Построение здания дифференциального исчисления в значительной своей части основано на теореме Лагранжа, являющейся первым образцом так называемых «теорем о среднем значении», которым мы, ввиду их крайне важного значения для всех разделов анализа, и в дальнейшем должны будем уделять достаточное внимание.

Теорема Лагранжа. *Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема внутри него, то существует такая внутренняя точка c этого отрезка, для которой*

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{A}$$

Доказательство. Очевидно, что функция

$$\varphi(x) = (b-a)[f(x)-f(a)] - (x-a)[f(b)-f(a)]$$

дифференцируема внутри отрезка $[a, b]$; так как $\varphi(a) = \varphi(b) = 0$, то либо $\varphi(x) = 0$ для всех $x \in [a, b]$, и тогда $\varphi'(x) = 0$ для всех таких x, либо же $\varphi(x)$ внутри отрезка принимает отличные от нуля значения, и тогда среди этих значений найдется либо наибольшее, либо наименьшее (так как функция $\varphi(x)$ непрерывна); пусть для определенности функция $\varphi(x)$ принимает наибольшее значение при $x = c$, где $a < c < b$; тогда необходимо $\varphi'(c) = 0$, ибо при $\varphi'(c) > 0$ мы имели бы для достаточно малых положительных h

$$\frac{\varphi(c+h) - \varphi(c)}{h} > 0,$$

откуда $\varphi(c+h) > \varphi(c)$, что противоречит определению точки c; в случае $\varphi'(c) < 0$ мы получаем аналогичное противоречие, давая h отрицательные значения; но из $\varphi'(c) = 0$ вытекает (A), и теорема доказана.

Очень часто ее формулируют в другой системе обозначений: если концы данного отрезка суть x и $x+\Delta x$, то говорят о существовании такого числа θ $(0<\theta<1)$, что $f'(x+\theta\,\Delta x)\,\Delta x = f(x+\Delta x) - f(x)$.

Для функций, обладающих производными высших порядков, теорема Лагранжа допускает важное обобщение, которое мы сейчас докажем.

Теорема. *Если функция* $y = f(x)$ *на отрезке* $[x,\ x+n\,\Delta x]$ *имеет непрерывную производную порядка* $n-1$, *а внутри отрезка — производную порядка* n [1]), *то существует такое число* θ $(0<\theta<1)$, *что*

$$\Delta^n y = f^{(n)}(x+\theta n\,\Delta x)\,(\Delta x)^n. \qquad \text{(B)}$$

Здесь $\Delta^n y$ означает так называемую «разность порядка n» данной функции $y = f(x)$; разность первого порядка есть просто приращение $\Delta y = f(x+\Delta x) - f(x)$; разность порядка $n+1$ определяется как (первая) разность от разности порядка n, так что, в частности,

$$\Delta^2 y = \Delta(\Delta y) =$$
$$= [f(x+\Delta x+\Delta x) - f(x+\Delta x)] - [f(x+\Delta x) - f(x)] =$$
$$= f(x+2\Delta x) - 2f(x+\Delta x) + f(x),$$
$$\Delta^3 y = f(x+3\Delta x) - 3f(x+2\Delta x) + 3f(x+\Delta x) - f(x)$$

и т. д. Вообще, как легко доказать методом полной индукции,

$$\Delta^n y = f(x+n\,\Delta x) - nf(x+[n-1]\,\Delta x) + \frac{n(n-1)}{2} f(x+[n-2]\,\Delta x) - \ldots$$
$$\ldots + (-1)^{n-1} nf(x+\Delta x) + (-1)^n f(x).$$

Доказательство. При $n=1$ доказываемая теорема совпадает с теоремой Лагранжа. Допустим поэто-

---

[1]) На концах отрезка достаточно существования односторонних производных порядка $n-1$.

му, что теорема верна для разностей порядка n, и покажем, что она верна и для разностей порядка $n+1$. Мы предполагаем, следовательно, что функция $y=f(x)$ имеет непрерывную производную порядка n на отрезке $[x, x+(n+1)\Delta x]$ и производную порядка $n+1$ внутри него; отсюда функция $\Delta y = f(x+\Delta x)-f(x)$ имеет производную порядка n на отрезке $[x, x+n\,\Delta x]$; мы можем, следовательно, применить нашу теорему к ее разности n-го порядка, что дает

$$\Delta^n(\Delta y)=(\Delta x)^n\{f^{(n)}(x+\Delta x+\theta_1 n\ \Delta x)-f^{(n)}(x+\theta_1 n\ \Delta x)\},$$

где $0<\theta_1<1$; но из наших предпосылок вытекает, что функция $f^{(n)}$ непрерывна на отрезке $[x+\theta_1 n\Delta x,\ x+ (1+\theta_1 n)\Delta x]$ и дифференцируема внутри него; поэтому, применяя к фигурной скобке предыдущего равенства теорему Лагранжа, находим

$$\Delta^{n+1}y=(\Delta x)^{n+1}f^{n+1}(x+\theta_1 n\ \Delta x+\theta_2\Delta x)=$$
$$=(\Delta x)^{n+1}f^{(n+1)}(x+\theta[n+1]\Delta x),$$

что и требовалось доказать.

Другое важное обобщение теоремы Лагранжа представляет собой замечательная «формула Коши». *Пусть мы имеем две функции $f_1(x)$ и $f_2(x)$, непрерывные на отрезке $[a, b]$ и дифференцируемые внутри него, причем $f_2'(x)$ нигде внутри этого отрезка не обращается в нуль. Тогда существует внутри $[a, b]$ такая точка c, что*

$$\frac{t_1(b)-f_1(a)}{f_2(b)-f_3(a)}=\frac{f_1'(c)}{f_2'(c)}. \qquad \text{(C)}$$

Доказательство этой общей формулы (при $f_2(x)=x$ переходящей в теорему Лагранжа) проводится совершенно так же, как и в случае теоремы Лагранжа; полагая

$$\varphi(x)=f_1(x)\,[f_2(b)-f_2(a)]-f_2(x)\,[f_1(b)-f_1(a)],$$

мы находим, что $\varphi(a)=\varphi(b)$; поэтому функция $\varphi(x)$ либо постоянна на отрезке $[a, b]$, либо принимает внутри этого отрезка наибольшее или наименьшее значение; отсюда в точности тем же путем, как в доказательстве теоремы Лагранжа, мы заключаем, что в некото-

рой точке c, $a < c < b$, мы должны иметь $\varphi'(c) = 0$, что равносильно утверждаемому равенству (C).

Заметим еще, что левая часть этого равенства не может оказаться лишенной смысла, так как при $f_2(a) = f_2(b)$ функция $f_2'(x)$ в силу теоремы Лагранжа должна была бы обращаться в нуль внутри $[a, b]$, что исключается предпосылками теоремы.

Повторное дифференцирование. Производные и дифференциалы высших порядков определяются, как вы знаете, индуктивно: производная порядка n получается дифференцированием производной порядка $n-1$, и аналогично обстоит дело для дифференциалов.

Производные и дифференциалы высших порядков функции $y=f(x)$ связаны между собой соотношениями

$$d^n y = f^{(n)}(x)\,dx^n \tag{5}$$

(где dx^n означает n-ю степень *первого* дифференциала независимой переменной x), которые доказываются также индуктивно: если x — независимая переменная, то $dx = \Delta x$ не зависит от x, а потому, дифференцируя соотношение (5), находим (штрих означает дифференцирование по x)

$$d^{n+1}y = d(d^n y) = (d^n y)'dx = f^{(n+1)}(x)\,dx^{n+1}.$$

Как мы видели выше, при $n = 1$ соотношение (5) остается верным и в том случае, когда x (а значит, и y) есть любая дифференцируемая функция независимой переменной t: $x = \varphi(t)$. Легко убедиться, что уже при $n = 2$ дело обстоит иначе: дифференциалы высших порядков не инвариантны относительно замены независимой переменной. В самом деле, будем обозначать второй дифференциал функции $y = f(x)$ через $(d^2y)_x$ или $(d^2y)_t$ в зависимости от того, считаем ли мы независимой переменной величину x или величину t. Тогда

$$(d^2y)_x = f''(x)\,dx^2,$$

между тем как

$$(d^2y)_t = \frac{d^2f[\varphi(t)]}{dt^2}dt^2 = \frac{d\{f'[\varphi(t)]\varphi'(t)\}}{dt}dt^2 =$$

$$= \{f''[\varphi(t)]\varphi'^2(t) + f'[\varphi(t)]\varphi''(t)\}\,dt^2 =$$

$$= f''(x)\,dx^2 + f'[\varphi(t)]\varphi''(t)\,dt^2.$$

Мы видим, что эти два выражения различны: второе содержит добавочный член $f'[\varphi(t)]\varphi''(t)dt^2$, которого нет в первом и который тождественно обращается в нуль только в том случае, когда x есть *линейная* функция от t: $\varphi(t) = at + b$; таким образом, второй дифференциал, в отличие от первого, инвариантен лишь относительно линейных преобразований независимой переменной.

Можно определять производные высших порядков и иначе, связывая их с разностями соответствующих порядков; подобно тому как первая производная является пределом отношения $\frac{\Delta y}{\Delta x}$ при $\Delta x \to 0$, можно показать, что

$$\lim_{\Delta x \to 0} \frac{\Delta^n y}{\Delta x^n} = f^{(n)}(x). \tag{6}$$

Если функция $f^{(n)}(x)$ непрерывна, то соотношение (6) является непосредственным следствием формулы (B), но важно установить, что справедливость соотношения (6) не зависит от этой предпосылки.

Так как соотношение (6) верно при $n = 1$, то надо только, предположив, что оно верно для некоторого числа n, убедиться, что оно должно оставаться верным и для числа $n + 1$ (причем, конечно, предполагается, что функция $y = f(x)$ имеет производную порядка $n + 1$). Пусть же соотношение (6) для данного n (и для любой n раз дифференцируемой функции y) справедливо. Так как функция

$$\Delta y = f(x + \Delta x) - f(x)$$

имеет производную порядка n, то мы можем применить к ней формулу (B); это дает (как мы уже видели, стр. 143)

$$\Delta^n(\Delta y) = \Delta x^n (\Delta y)^{(n)}_{x+\theta n\,\Delta x} =$$

$$= \Delta x^n \{f^{(n)}(x + \theta n\,\Delta x + \Delta x) - f^{(n)}(x + \theta n\,\Delta x)\}, \tag{7}$$

где $0 < \theta < 1$. Но так как в силу существования $f^{(n+1)}(x)$

$$f^{(n)}(x + \theta n\,\Delta x + \Delta x) - f^{(n)}(x) = (1 + n\theta)\,\Delta x\{f^{(n+1)}(x) + \alpha_1\},$$

$$f^{(n)}(x + \theta n\,\Delta x) - f^{(n)}(x) = n\theta\,\Delta x\{f^{(n+1)}(x) + \alpha_2\},$$

где $\alpha_1 \to 0$ и $\alpha_2 \to 0$ при $\Delta x \to 0$, то

$$f^{(n)}(x+\theta n\,\Delta x+\Delta x)-f^{(n)}(x+\theta n\,\Delta x)=$$
$$=\Delta x f^{(n+1)}(x)+\alpha_1(1+n\theta)\,\Delta x-\alpha_2 n\theta\,\Delta x,$$

вследствие чего соотношение (7) дает

$$\frac{\Delta^{n+1}y}{\Delta x^{n+1}}=f^{(n+1)}(x)+\alpha_1(1+n\theta)-\alpha_2 n\theta,$$

откуда в пределе при $\Delta x \to 0$ получается

$$\lim_{\Delta x\to 0}\frac{\Delta^{n+1}y}{\Delta x^{n+1}}=f^{(n+1)}(x),$$

что и требовалось установить.

Пределы отношений бесконечно малых и бесконечно больших. Среди приложений формул Лагранжа и Коши рассмотрим в первую очередь важный вопрос, фигурирующий обычно в курсах анализа под совершенно нелепым наименованием «раскрытия неопределенностей»[1]. По существу здесь речь идет об использовании методов дифференциального исчисления для отыскания пределов отношений двух бесконечно малых или двух бесконечно больших величин.

Пусть $f_1(a)=f_2(a)=0$, причем в некоторой окрестности точки a обе функции дифференцируемы, и $f_2'(x)\neq 0$ при $x\neq a$. Так как при $f_1(a)=f_2(a)=0$ формула Коши (C) дает

$$\frac{f_1(a+h)}{f_2(a+h)}=\frac{f_1'(a+\theta h)}{f_2'(a+\theta h)}\qquad(0<\theta<1),$$

то мы можем утверждать (правило Лопиталя): *если* $f_1(a)=f_2(a)=0$ *и отношение* $\dfrac{f_1'(x)}{f_2'(x)}$ *стремится к*

[1]) Еще более нелепо наименование «нахождение истинных значений неопределенных выражений»; не говоря уже о том, что никаких «неопределенных выражений» анализ вообще не содержит, здесь «истинным значением» функции в данной точке называется число, вообще говоря отличное от ее фактического значения.

некоторому пределу при $x \to a$, *то к тому же пределу стремится и* $\frac{f_1(x)}{f_2(x)}$.

Если $f_1'(a) = f_2'(a) = 0$, то, повторяя это рассуждение, мы находим (предполагая, разумеется, существование вторых производных функций $f_1(x)$ и $f_2(x)$ в некоторой окрестности точки a): при $f_1(a) = f_2(a) = f_1'(a) = f_2'(a) = 0$ из соотношения $\lim\limits_{x\to a} \frac{f_1''(x)}{f_2''(x)} = l$ вытекает, что и $\lim\limits_{x\to a} \frac{f_1(x)}{f_2(x)} = l$; вообще, если

$$f_1(a) = f_1'(a) = \ldots = f_1^{(n-1)}(a) = f_2(a) = \\ = f_2'(a) = \ldots = f_2^{(n-1)}(a) = 0$$

и если функции $f_1(x)$ и $f_2(x)$ в некоторой окрестности точки a имеют производные порядка n, то из $\frac{f_1^{(n)}(x)}{f_2^{(n)}(x)} \to l$ при $x \to a$ вытекает, что и $\frac{f_1(x)}{f_2(x)} \to l$ при $x \to a$. Если $f_1^{(n)}(x)$ и $f_2^{(n)}(x)$ непрерывны в точке a и $f_2^{(n)}(a) \neq 0$, то отсюда следует, что

$$\lim_{x\to a} \frac{f_1(x)}{f_2(x)} = \frac{f_1^{(n)}(a)}{f_2^{(n)}(a)}. \tag{8}$$

Правило Лопиталя является весьма мощным орудием вычисления пределов и во многих случаях позволяет с большой легкостью находить эти пределы там, где применение элементарных методов вызвало бы большие затруднения. При $f_1(x) = \operatorname{tg} x - \sin x$, $f_2(x) = x^3$ мы имеем $f_1(0) = f_1'(0) = f_1''(0) = 0$, $f_1'''(0) = 3$, $f_2(0) = f_2'(0) = f_2''(0) = 0$, $f_2'''(0) = 6$, вследствие чего по формуле (8)

$$\lim_{x\to 0} \frac{\operatorname{tg} x - \sin x}{x^3} = \frac{f_1'''(0)}{f_2'''(0)} = \frac{1}{2}.$$

Правило Лопиталя остается в силе и в случае $a = \pm\infty$. В самом деле,

$$\lim_{x \to +\infty} \frac{f_1(x)}{f_2(x)} = \lim_{y \to +0} \frac{f_1\left(\frac{1}{y}\right)}{f_1\left(\frac{1}{y}\right)};$$

в силу же доказанного нами правила Лопиталя последний предел совпадает с

$$\lim_{y \to +0} \frac{-\frac{1}{y^2} f_1'\left(\frac{1}{y}\right)}{-\frac{1}{y^2} f_2'\left(\frac{1}{y}\right)} = \lim_{x \to +\infty} \frac{f_1'(x)}{f_2'(x)}$$

в случае, когда этот предел существует; конечно, необходимой предпосылкой правила здесь служит условие

$$\lim_{x \to +\infty} f_1(x) = \lim_{x \to +\infty} f_2(x) = 0.$$

Более того, правило Лопиталя может быть применяемо и к отысканию предела отношения бесконечно больших величин, хотя здесь доказательство его несколько сложнее. Пусть для определенности $f_1(x) \to +\infty$ и $f_2(x) \to +\infty$ при $x \to a$; кроме того, мы предположим, как и прежде, что $f_2'(x)$ не обращается в нуль в некоторой окрестности точки a. Возьмем в этой окрестности две любые точки x и α так, чтобы x лежало между a

a x ξ α

Рис. 22.

и α; пусть для определенности $a < x < \alpha$ (рис. 22); в силу формулы Коши (C) мы имеем

$$\frac{f_1(x) - f_1(\alpha)}{f_2(x) - f_2(\alpha)} = \frac{f_1'(\xi)}{f_2'(\xi)}, \tag{9}$$

где ξ — некоторая внутренняя точка отрезка $[x, \alpha]$.

Так как, с другой стороны,

$$\frac{f_1(x)-f_1(\alpha)}{f_2(x)-f_2(\alpha)}=\frac{f_1(x)}{f_2(x)}\,\frac{1-\dfrac{f_1(\alpha)}{f_1(x)}}{1-\dfrac{f_2(\alpha)}{f_2(x)}},$$

то равенство (9) дает нам

$$\frac{f_1(x)}{f_2(x)}=\frac{f_1'(\xi)}{f_2'(\xi)}\,\frac{1-\dfrac{f_2(\alpha)}{f_2(x)}}{1-\dfrac{f_1(\alpha)}{f_1(x)}}. \tag{10}$$

Теперь мы распорядимся выбором точек x и α более детально. Если, как мы предполагаем, существует

$$\lim_{x\to a}\frac{f_1'(x)}{f_2'(x)}=l,$$

то при сколь угодно малом $\varepsilon>0$ найдется такая окрестность U точки a, что для всех $x\in U$

$$l-\varepsilon<\frac{f_1'(x)}{f_2'(x)}<l+\varepsilon;$$

пусть α (а значит, и x) принадлежит этой окрестности; тогда и $\xi\in U$, а потому $l-\varepsilon<\dfrac{f_1'(\xi)}{f_2'(\xi)}<l+\varepsilon$; оставим теперь α неизменным, а x будем неограниченно приближать к a; так как при этом $f_1(x)\to+\infty$ и $f_2(x)\to\to+\infty$, то второй множитель правой части формулы (10) стремится к единице, так что для всех x в некоторой окрестности V точки a он будет заключен между $1-\varepsilon$ и $1+\varepsilon$, вследствие чего для всех $x\in V$ формула (10) дает

$$(l-\varepsilon)(1-\varepsilon)<\frac{f_1(x)}{f_2(x)}<(l+\varepsilon)(1+\varepsilon);$$

так как ε сколь угодно мало, то

$$\lim_{x\to a}\frac{f_1(x)}{f_2(x)}=l,$$

что и требовалось доказать.

Это правило, подобно правилу для отношения двух бесконечно малых, остается в силе и при $a=\pm\infty$.

Примеры: 1) $f_1(x) = \ln x$, $f_2(x) = \frac{1}{x}$ при $x \to +0$. Здесь $f_1(x) \to -\infty$, $f_2(x) \to +\infty$, $\frac{f_1'(x)}{f_2'(x)} = -x \to 0$; отсюда $\frac{f_1(x)}{f_2(x)} = x \ln x \to 0$ при $x \to +0$. 2) При $f_2(x) = x \to +\infty$ мы имеем $f_1(x) = \ln x \to +\infty$, причем $\frac{f_1'(x)}{f_2'(x)} = \frac{1}{x} \to 0$; отсюда $\frac{f_1(x)}{f_2(x)} = \frac{\ln x}{x} \to 0$ при $x \to +\infty$; полагая $x = e^y$, находим: $ye^{-y} \to 0$ при $y \to +\infty$.

Формула Тэйлора. Теперь мы должны остановиться на формуле Тэйлора, служащей, как вам, вероятно, известно, одним из важнейших орудий исследования как в самом анализе, так и в его приложениях. Эта формула, подробно рассматриваемая во всех курсах анализа, требует нашего внимания потому, что при обычных изложениях ее вывода идейная сторона дела остается совершенно в тени, уступая место ничем не оживленному формализму, вследствие чего для многих учащихся в дальнейшем то важное значение, которое получает эта формула, является неожиданным, а подчас и навсегда остается загадкой.

Как мы уже не раз отмечали, основное соотношение

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = f'(a)$$

может быть записано в эквивалентном виде

$$f(a+h) = f(a) + hf'(a) + \alpha h, \qquad (11)$$

где $\alpha \to 0$ при $h \to 0$; здесь αh есть, следовательно, бесконечно малая высшего порядка относительно h (отношение ее к h стремится к нулю при $h \to 0$).

Условимся вообще обозначать через $o(x)$ всякую величину, отношение которой к величине x стремится к нулю в данном процессе изменения; тогда произведение αh мы можем записать в виде $o(h)$; так как для всякого $o(h)$ по определению

$$\frac{o(h)}{h} = \alpha \to 0 \qquad (h \to 0),$$

то и, обратно, всякое $o(h)$ может быть представлено

в виде αh, где $\alpha \to 0$ при $h \to 0$; поэтому соотношение (11) может быть записано в виде

$$f(a+h) = f(a) + hf'(a) + o(h);$$

как мы видим, эта формула имеет место во всех случаях, когда существует $f'(a)$; мы уже не раз имели случай применять ее, и непосредственно ясно, чему она обязана этой своей плодотворностью: если h столь мало, что величиной $o(h)$ можно пренебрегать, то приближенная формула

$$f(a+h) \approx f(a) + hf'(a)$$

позволяет нам в целях исследования заменять ту сложную, вообще говоря, функцию от h, какую представляет собой $f(a+h)$, линейной функцией, стоящей в правой части; ясно, какую это дает существенную выгоду.

Теперь будет совершенно понятным, если мы пожелаем продолжить это полезное мероприятие несколько дальше. Если мы хотим получить лучшую точность, то мы иногда не можем довольствоваться приближенными формулами, погрешность которых мала сравнительно с h; может, например, случиться, что мы должны еще учитывать величины порядка h^2, а пренебрегать можем лишь величиной $o(h^2)$, т. е. бесконечно малой сравнительно с h^2; тогда мы, естественно, ставим вопрос об отыскании такой функции второй степени $a_0+a_1h+a_2h^2$, чтобы при малых h имело место соотношение

$$f(a+h) = a_0 + a_1h + a_2h^2 + o(h^2);$$

вообще, если мы хотим учитывать величины порядка h^n, но соглашаемся пренебрегать величиной $o(h^n)$, то, мы, естественно, ищем такой многочлен $a_0 + a_1h + \ldots + a_nh^n = P_n(h)$, чтобы имело место соотношение

$$f(a+h) = P_n(h) + o(h^n);$$

если эту задачу удается решить, то мы получаем возможность во всех вопросах, где можно пренебрегать величиной $o(h^n)$, вместо сложной, вообще говоря, функции $f(a+h)$ пользоваться простым многочленом n-й степени.

В случае $n = 1$ мы получили решение поставленной задачи, предполагая, что функция $f(x)$ имеет первую

производную при $x = a$; в общем случае мы, естественно, будем исходить из предположения, что существует $f^{(n)}(a)$; очевидно, это требует от функции $f(x)$ существования производных всех низших порядков в некоторой окрестности точки a.

Знаменитая теорема Тэйлора в основном как раз и дает решение поставленной нами задачи, т. е. доказывает существование искомого многочлена $P_n(h)$ и дает выражение для его коэффициентов. В силу этой теоремы в случае, если $f^{(n)}(a)$ существует, многочлен $P_n(h)$ однозначно определяется формулой

$$P_n(h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots + \frac{h^n}{n!}f^{(n)}(a).$$

Иначе говоря, в силу этой теоремы

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots$$
$$\ldots + \frac{h^n}{n!}f^{(n)}(a) + o(h^n). \qquad (12)$$

Положим

$$f(a+h) - P_n(h) = \varphi(h);$$

тогда для доказательства нашего утверждения достаточно убедиться, что

$$\frac{\varphi(h)}{h^n} \to 0 \text{ при } h \to 0. \qquad (12')$$

С этой целью заметим, что

$$\varphi(h) = f(a+h) - f(a) - hf'(a) - \ldots$$
$$\ldots - \frac{h^n}{n!}f^{(n)}(a),$$

$$\varphi'(h) = f'(a+h) - f'(a) - hf''(a) - \ldots$$
$$\ldots - \frac{h^{n-1}}{(n-1)!}f^{(n)}(a),$$

. . . .

$$\varphi^{(n-2)}(h) = f^{(n-2)}(a+h) - f^{(n-2)}(a) - hf^{(n-1)}(a) -$$
$$- \frac{h^2}{2}f^{(n)}(a),$$

$$\varphi^{(n-1)}(h) = f^{(n-1)}(a+h) - f^{(n-1)}(a) - hf^{(n)}(a),$$

откуда $\varphi(0) = \varphi'(0) = \ldots = \varphi^{(n-2)}(0) = 0$. Так как производные функции h^n до порядка $n-2$ включительно также обращаются в нуль при $h = 0$, а производная порядка $n-1$ равна $n!h$, то по правилу Лопиталя

$$\lim_{h\to 0} \frac{\varphi(h)}{h^n} = \lim_{h\to 0} \frac{\varphi^{(n-1)}(h)}{n!\,h},$$

если только предел правой части существует; но

$$\lim_{h\to 0} \frac{\varphi^{(n-1)}(h)}{n!h} = \frac{1}{n!} \lim_{h\to 0} \left\{ \frac{f^{(n-1)}(a+h) - f^{(n-1)}(a)}{h} - \right.$$
$$\left. - f^{(n)}(a) \right\} = 0,$$

так как $f^{(n)}(a)$ по предположению существует; этим доказано соотношение (12′), и тем самым закончен вывод формулы Тэйлора (12).

В дальнейшем представляет, конечно, интерес найти удобные выражения для «остаточного члена» $o(h^n)$ формулы (12), чтобы в нужных случаях оценивать его с возможной точностью.

Мы здесь напомним наиболее употребительные формы этих выражений, не приводя их выводов, которые вы можете найти в любом полном курсе анализа. Обозначим через $R_n(h)$ величину, имеющую в формуле (12) анонимное обозначение $o(h^n)$. В силу самой формулы Тэйлора мы при существовании $f^{(n+1)}(a)$ имеем

$$\lim_{h\to 0} \frac{R_n(h)}{h^{n+1}} = \frac{f^{(n+1)}(a)}{(n+1)!};$$

если мы допустим, что $f^{(n+1)}(x)$ существует не только при $x = a$, но и во всех точках отрезка $[a, a+h]$, то мы можем получить для $R_n(h)$ известную «форму Шлёмильха»:

$$R_n(h) = \frac{h^{n+1}(1-\theta)^{n+1-p}}{n!\,p} f^{(n+1)}(a+\theta h),$$

где $0 < \theta < 1$, а p — *произвольное* положительное число $\leqslant n + 1$. Из этой весьма общей формы при специальных выборах числа p получается ряд различных особых форм; наиболее употребительные из них получаются:

1) при $p = n + 1$ «форма Лагранжа»:

$$R_n(h) = \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(a + \theta h);$$

2) при $p = 1$ «форма Коши»:

$$R_n(h) = \frac{h^{n+1}(1-\theta)^n}{n!} f^{(n+1)}(a + \theta h).$$

В этих специальных формах, как и в общей форме Шлёмильха, предполагается, что $f^{(n+1)}(x)$ существует в каждой точке отрезка $[a, a + h]$.

Вы знаете, что *формулой Маклорена* называют (ничем особо не замечательный) частный случай формулы Тэйлора при $a = 0$:

$$f(h) = f(0) + \frac{h}{1} f'(0) + \ldots + \frac{h^n}{n!} f^{(n)}(0) + o(h^n).$$

Экстремальные значения. Важной областью применения формулы Тэйлора в элементах дифференциального исчисления служит теория максимальных и минимальных значений, с основами которой вы, конечно, знакомы.

Если разыскивается, скажем, точка отрезка $[a, b]$, в которой функция $y = f(x)$, дифференцируемая на этом отрезке, принимает свое наибольшее значение, то этот «абсолютный максимум» может, очевидно, осуществиться либо в одном из концов отрезка $[a, b]$, либо в некоторой внутренней его точке c; в последнем случае точка c есть вместе с тем точка «относительного максимума»; это значит, что $f(c) \geqslant f(x)$ для всех x, принадлежащих некоторой окрестности точки c. Таким образом, отыскание абсолютного максимума функции на данном отрезке сводится к отысканию всех ее отно-

сительных максимумов на этом отрезке; именно к этой задаче и применяются методы дифференциального исчисления.

Вы знаете, конечно, что необходимым условием для того, чтобы точка x $(a < x < b)$ давала относительный максимум (или минимум) дифференцируемой функции $f(x)$, служит равенство

$$f'(x) = 0. \tag{13}$$

Таким образом, первым шагом к решению поставленной задачи является отыскание всех вещественных корней уравнения (13) — так называемых «критических значений» переменной x; после того как это сделано, каждое отдельное критическое значение подвергается особому исследованию, которое должно установить, имеет ли место в данной точке максимум или минимум или нет ни того, ни другого; вот на этом-то исследовании мы теперь немного остановимся.

Пусть α — одно из критических значений x для функции $f(x)$, т. е. $a < \alpha < b$ и $f'(\alpha) = 0$; допустим для общности, что $f^{(i)}(\alpha) = 0$ при $1 \leqslant i < n$, но $f^{(n)}(\alpha) \neq 0$. В таком случае формула Тэйлора (12) дает

$$f(\alpha + h) - f(\alpha) = h^n \frac{f^{(n)}(\alpha)}{n!} + o(h^n); \tag{14}$$

очевидно, что знак правой части, второе слагаемое которой бесконечно мало сравнительно с первым, при достаточно малых $|h|$ совпадает со знаком ее первого слагаемого, т. е. со знаком произведения $h^n f^{(n)}(\alpha)$. Если число n нечетное, то h^n, а значит и это произведение, меняет знак при изменении знака величины h; в силу (14) в этом случае и знак разности $f(\alpha + h) - f(\alpha)$ будет различным для приращений h разного знака, а это, очевидно, означает, что при $x = \alpha$ функция $f(x)$ не может иметь ни максимума, ни минимума. Пусть теперь n — четное число, так что $h^n > 0$ при любом $h \neq 0$; тогда знак разности $f(\alpha + h) - f(\alpha)$ для всех достаточно малых $|h|$ совпадает, очевидно, со знаком величины $f^{(n)}(\alpha)$, уже не зависящей от h; в случае

$f^{(n)}(\alpha) > 0$ это означает, что для всех достаточно малых $|h|$

$$f(\alpha + h) > f(\alpha),$$

т. е. что в точке α функция $f(x)$ имеет относительный минимум; точно так же мы убеждаемся, что в случае $f^{(n)}(\alpha) < 0$ функция $f(x)$ имеет относительный максимум при $x = \alpha$. Таким образом, мы приходим к следующему правилу: *пусть среди производных функции $f(x)$ при $x = \alpha$ первая не обращающаяся в нуль имеет порядок n; тогда $f(x)$ при четном n имеет в точке $x = \alpha$ максимум, если $f^{(n)}(\alpha) < 0$, и минимум, если $f^{(n)}(\alpha) > 0$; при нечетном же n функция $f(x)$ не имеет в точке α ни максимума, ни минимума.* Вы видите, что этим правилом вопрос о природе каждой критической точки решается до конца, если только функция дифференцируема в этой точке достаточное число раз и если не все ее производные обращаются в нуль в этой точке.

Частное дифференцирование. Теперь мы должны обратиться к функциям нескольких переменных; для простоты мы ограничимся при этом рассмотрением случая двух независимых переменных x и y.

Вы знаете, конечно, что *частная производная* $f'_x(x, y)$ или $\frac{\partial f(x, y)}{\partial x}$ функции $f(x, y)$ по переменной x определяется как

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x},$$

т. е. как обыкновенная производная той функции от x, в которую переходит функция $f(x, y)$ после придания переменной y какого-либо постоянного значения; аналогично определяют $\frac{\partial f}{\partial y} = f'_y(x, y)$; каждая из этих частных производных есть функция тех же двух переменных x и y, от которых зависит и $f(x, y)$. Ради удобства и наглядности мы будем и здесь часто слова «при $x = a$, $y = b$» заменять выражением «в точке (a, b)», имея в виду точку плоскости xy с декартовыми координатами $x = a$, $y = b$.

Прежде всего надо определить понятие дифференцируемости данной функции $z = f(x, y)$ в данной точке

(a, b). В случае функции одной переменной существование производной, конечно, равносильно существованию дифференциала; при этом, как мы видели, дифференциал dy может быть определен для функции $y=f(x)$ как такая линейная функция приращения Δx, которая при $\Delta x \to 0$ отличается от приращения Δy функции y на бесконечно малую высшего порядка.

Посмотрим теперь, как обстоит дело в случае функции двух переменных $z = f(x, y)$; условимся и здесь называть дифференциалом dz функции z такую линейную комбинацию $A\,\Delta x + B\,\Delta y$ приращений Δx и Δy, которая при $\Delta x \to 0$, $\Delta y \to 0$ отличается от приращения Δz на бесконечно малую высшего порядка; в случае одномерном мы могли при этом (предполагая, что $f'(x) \neq 0$) принять за основную бесконечно малую любую из величин Δx, Δy, dy (порядки малости которых совпадают между собой); в двумерном случае удобно (хотя отнюдь не обязательно) принять за основную бесконечно малую величину $\rho = \sqrt{\Delta x^2 + \Delta y^2}$, выражающую собой расстояние «смещенной» точки $(x + \Delta x, y + \Delta y)$ от «первоначальной» (x, y); если бесконечно малые Δx и Δy имеют одинаковый порядок малости, то тот же порядок имеет, конечно, и величина ρ. Таким образом, выражение $A\,\Delta x + B\,\Delta y$ есть дифференциал dz функции $z = f(x, y)$, если при $\Delta x \to 0$, $\Delta y \to 0$

$$\Delta z - (A\,\Delta x + B\,\Delta y) = o(\rho). \qquad (15)$$

Покажем теперь, что если dz существует, то существуют и частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, причем

$$A = \frac{\partial z}{\partial x}, \qquad B = \frac{\partial z}{\partial y},$$

так что

$$dz = \frac{\partial z}{\partial x}\,\Delta x + \frac{\partial z}{\partial y}\,\Delta y. \qquad (16)$$

В самом деле, прежде всего положим в формуле (15) $\Delta y = 0$ (по нашему условию эта формула должна иметь место при *любом* законе стремления величин

Δx и Δy к нулю); это дает

$$(\Delta z)_{\Delta y=0} = A\,\Delta x + o(\Delta x),$$

откуда

$$A = \lim_{\Delta x \to 0} \frac{(\Delta z)_{\Delta y=0}}{\Delta x} = \frac{\partial z}{\partial x};$$

совершенно аналогично устанавливается и то, что $B = \frac{\partial z}{\partial y}$. При этом, очевидно, самое существование частных производных $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ вытекает из наших рассуждений и является, таким образом, следствием существования дифференциала dz.

До сих пор все протекает, как вы видите, в полной аналогии с одномерным случаем.

Однако, в противоположность одномерному случаю, существование в данной точке частных производных $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ еще не гарантирует существования в ней дифференциала dz; так, для функции

$$z = \begin{cases} \dfrac{2xy}{\sqrt{x^2+y^2}} & \text{при } x^2+y^2 \neq 0, \\ 0 & \text{при } x = y = 0 \end{cases}$$

в точке $(0, 0)$ мы имеем, очевидно,

$$(\Delta z)_{\Delta x=0} = (\Delta z)_{\Delta y=0} = 0,$$

вследствие чего

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0;$$

если бы дифференциал dz существовал при $x = y = 0$, то он должен был бы поэтому в силу (16) тождественно равняться нулю, т. е. мы должны были бы в силу (15) иметь $\Delta z = o(\rho)$; но это неверно, так как, например, при $\Delta x = \Delta y \neq 0$, очевидно,

$$\Delta z = \sqrt{2}\Delta x, \quad \rho = \sqrt{2}|\Delta x|, \quad |\Delta z| = \rho.$$

Заметим, впрочем, что если $\frac{\partial f}{\partial x}$ и $\frac{\partial f}{\partial y}$ *непрерывны* в точке (x, y), то в этой точке выражение (16) есть дифференциал функции $z = f(x, y)$. В самом деле,

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y) =$$
$$= [f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)] +$$
$$+ [f(x, y + \Delta y) - f(x, y)].$$

Первая из разностей правой части в силу теоремы Лагранжа имеет вид

$$\Delta x f'_x (x + \theta \Delta x, \; y + \Delta y) \qquad (0 < \theta < 1),$$

а так как в силу предположенной непрерывности функции $f'_x(x, y)$ в точке (x, y) при $\Delta x \to 0$, $\Delta y \to 0$

$$f'_x (x + \theta \Delta x, \; y + \Delta y) - f'_x (x, y) \to 0,$$

то указанная первая разность отличается от $f'_x(x, y)\, \Delta x$ на величину $o(\Delta x) = o(\rho)$:

$$f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) = f'_x (x, y)\, \Delta x + o(\rho);$$

а так как, с другой стороны, очевидно, по самому определению частных производных

$$f(x, y + \Delta y) - f(x, y) = f'_y (x, y)\, \Delta y + o(\Delta y) =$$
$$= f'_y (x, y)\, \Delta y + o(\rho),$$

то

$$\Delta z = f'_x (x, y)\, \Delta x + f'_y (x, y)\, \Delta y + o(\rho),$$

что и требовалось доказать. Как показывает наше рассуждение, достаточно предположить даже непрерывность одной из двух частных производных в точке (x, y).

Как бы то ни было, в общем случае существование частных производных возможно и без существования дифференциала. Поэтому при определении дифференцируемости функции двух переменных перед нами стоит выбор: опираться ли на существование частных производных или на существование дифференциала?

Общепринято *называть функцию дифференцируемой в данной точке, если она в этой точке имеет дифференциал*; такое определение, требующее от дифференцируемой функции несколько больше, чем существование частных производных, удобно потому, что, как показывает дальнейшее развитие теории, только функции, обладающие дифференциалом, обнаруживают в своих свойствах сколько-нибудь полное сходство с дифференцируемыми функциями одной независимой переменной.

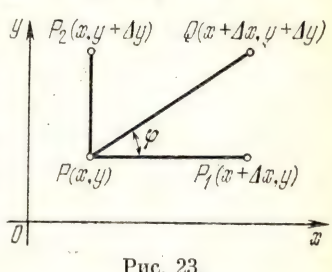


Рис. 23.

Для получения частной производной $\frac{\partial z}{\partial x}$ мы даем приращение одной только переменной x, оставляя y неизменным; говоря геометрически, мы *смещаем* точку $P(x, y)$ параллельно Ox в точку $P_1(x+\Delta x, y)$ (рис. 23); частная производная $\frac{\partial z}{\partial x}$ определяется как предел отношения

$$\frac{f(P_1)-f(P)}{PP_1}$$

при $P_1 \to P$; здесь $PP_1 = \Delta x$ есть расстояние между точками P и P_1, $f(P) = f(x, y)$, $f(P_1) = f(x+\Delta x, y)$. Подобным же образом,

$$\frac{\partial z}{\partial y} = \lim_{P_2 \to P} \frac{f(P_2)-f(P)}{PP_2},$$

где P_2 — точка с координатами $(x, y+\Delta y)$, которую мы неограниченно приближаем к точке P вдоль прямой, параллельной Oy. Можно, таким образом, сказать, что $\frac{\partial z}{\partial x}$ есть производная функции z в точке P «по направлению Ox», а $\frac{\partial z}{\partial y}$ — производная «по направлению Oy». Но ясно, что совершенно аналогичным образом мы можем определить частную производную функции $z = f(x, y)$ и по любому другому направлению, составляющему произвольный угол φ с осью Ox; для этого на-

до только сместить точку P в точку Q так, чтобы вектор $\overline{PQ}$ составлял угол φ с направлением Ox, и приближать затем точку Q к точке P вдоль этого вектора. Величину

$$\lim_{Q \to P} \frac{f(Q) - f(P)}{PQ},$$

если она существует, естественно называть *производной функции z по данному, характеризуемому углом φ, направлению*. Очевидно, что задание такого направления равносильно установлению определенной линейной связи ($\Delta y = \operatorname{tg} \varphi \, \Delta x$) между приращениями Δx и Δy при их одновременном стремлении к нулю.

Можно условиться обозначать, например, через $D_\varphi(z)$ производную функции z по направлению, составляющему угол φ с направлением оси Ox, так что

$$\frac{\partial z}{\partial x} = D_0(z), \quad \frac{\partial z}{\partial y} = D_{\pi/2}(z).$$

Докажем теперь, что *функция $z = f(x, y)$, дифференцируемая в точке (x, y), имеет в этой точке производную по любому направлению, причем*

$$D_\varphi(z) = \frac{\partial z}{\partial x} \cos \varphi + \frac{\partial z}{\partial y} \sin \varphi.$$

В самом деле, так как при смещении точки (x, y) по направлению φ на расстояние ρ

$$\Delta z = \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y + o(\rho),$$

причем

$$\Delta x = \rho \cos \varphi, \quad \Delta y = \rho \sin \varphi,$$

то

$$\frac{\Delta z}{\rho} = \frac{\partial z}{\partial x} \cos \varphi + \frac{\partial z}{\partial y} \sin \varphi + o(1),$$

т. е.

$$\lim_{\rho \to 0} \frac{\Delta z}{\rho} = D_\varphi(z) = \frac{\partial z}{\partial x} \cos \varphi + \frac{\partial z}{\partial y} \sin \varphi,$$

что и требовалось доказать.

Неявные функции. Дальнейшее развитие учения о частном дифференцировании совершается в достаточно полной аналогии с дифференциальным исчислением одномерного случая; за недостатком времени мы поэтому его касаться не будем; зато нам необходимо остановиться еще на некоторых важных задачах одномерного случая, для решения которых применяется частное дифференцирование. Сюда относятся прежде всего теоремы о существовании и дифференцировании «неявных» функций (этот общепринятый термин, безусловно, неудачен; правильно было бы говорить о «неявно заданных» или «неявно определенных» функциях, так как речь идет не об особом классе, а именно об особом способе задания функций).

Вам, конечно, приходилось много раз слышать о том, что уравнение

$$F(x, y) = 0 \tag{17}$$

определяет величину y как «неявную» функцию величины x. Под этим имеют в виду существование такой функции $y = f(x)$, что

$$F(x, f(x)) = 0 \tag{18}$$

тождественно. Само собой разумеется, что условия, обеспечивающие существование такой функции, свойства ее, а также и область значений x, для которых выполняется соотношение (18), подлежат специальному исследованию. Мы здесь докажем в этом направлении только основную теорему; основанное на ней дальнейшее развитие теории «неявных» функций вы найдете в любом полном курсе анализа; оно уже не представляет принципиальной новизны.

Теорема. *Пусть функция* $F(x, y)$: 1) *непрерывна в некоторой окрестности точки* (x_0, y_0), 2) *дифференцируема в точке* (x_0, y_0) *и* 3) *обращается в этой точке в нуль, в то время как* $F'_y(x_0, y_0) \neq 0$. *Тогда существует функция* $y = f(x)$, *для которой* $f(x_0) = y_0$ *и* (18) *тождественно выполняется в некоторой окрестности точки* x_0. *Эта функция* $f(x)$ *дифференцируема в*

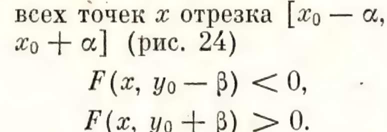
точке x_0, *и*

$$f'(x_0) = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)}.$$

Доказательство. Для определенности допустим, что $F'_y(x_0, y_0) > 0$; тогда при достаточно малом $\beta > 0$ мы будем иметь

$$F(x_0, y_0 - \beta) < 0, \quad F(x_0, y_0 + \beta) > 0,$$

причем функция $F(x, y)$ в обеих этих точках непрерывна; поэтому, если $\alpha > 0$ достаточно мало, то для всех точек x отрезка $[x_0 - \alpha, x_0 + \alpha]$ (рис. 24)



$$F(x, y_0 - \beta) < 0,$$
$$F(x, y_0 + \beta) > 0.$$

Рис. 24.

В силу той же непрерывности функции F для каждого такого значения x найдется поэтому, на основании теоремы 3 лекции III (стр. 77), такое значение y ($|y - y_0| < \beta$), для которого $F(x, y) = 0$. Это значение y (или одно из них, если их несколько), зависящее от x, мы будем обозначать через $f(x)$, так что

$$F(x, f(x)) = 0 \qquad (18)$$

при $x_0 - \alpha \leqslant x \leqslant x_0 + \alpha$; при $x = x_0$ функция $f(x)$, очевидно, имеет лишь одно возможное значение y_0, так что $f(x_0) = y_0$.

Нам остается доказать существование $f'(x_0)$ и найти величину этой производной. Пусть $|\Delta x| \leqslant \alpha$, $\Delta y = f(x_0 + \Delta x) - f(x_0)$; тогда в силу (18) $F(x_0 + \Delta x, y_0 + \Delta y) = F(x_0 + \Delta x, f(x_0 + \Delta x)) = 0$, откуда в силу дифференцируемости функции $F(x, y)$ в

точке (x_0, y_0)

$$0 = F(x_0 + \Delta x, y_0 + \Delta y) - F(x_0, y_0) = \Delta F =$$
$$= F'_y(x_0, y_0)\,\Delta y + F'_x(x_0, y_0)\,\Delta x + o(\rho),$$

где

$$\rho = \sqrt{\Delta x^2 + \Delta y^2} \leqslant |\Delta x| + |\Delta y|.$$

Отсюда

$$F'_y(x_0, y_0)\,\Delta y + F'_x(x_0, y_0)\,\Delta x = \lambda\,(|\Delta x| + |\Delta y|),$$

где $\lambda \to 0$ при $\rho \to 0$; иначе,

$$[F'_y(x_0, y_0) \pm \lambda]\,\Delta y + [F'_x(x_0, y_0) \pm \lambda]\,\Delta x = 0;$$

так как $\gamma = F'_y(x_0, y_0) > 0$, то при достаточно малых α и β (а значит, и достаточно малых $|\Delta x|$, $|\Delta y|$ и ρ) мы имеем

$$|\lambda| < \frac{1}{2}\gamma, \quad F'_y(x_0, y_0) \pm \lambda = \gamma \pm \lambda > \frac{1}{2}\gamma,$$

а следовательно,

$$\frac{\Delta y}{\Delta x} = -\frac{F'_x(x_0, y_0) \pm \lambda}{F'_y(x_0, y_0) \pm \lambda}, \quad \left|\frac{\Delta y}{\Delta x}\right| < \frac{|F'_x(x_0, y_0)| + \frac{1}{2}\gamma}{\frac{1}{2}\gamma};$$

отсюда прежде всего видно, что при $\Delta x \to 0$ и $\Delta y \to 0$, а потому и $\rho \to 0$, $\lambda \to 0$; значит,

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)},$$

что и требовалось доказать.

Заметим еще, что если F'_y существует и $\neq 0$ не только в точке (x_0, y_0), но и в некоторой ее окрестности, то найденное нами решение уравнения (17) является единственным; в самом деле, при $y_1 < y_2$, $F(x, y_1) = 0$, $F(x, y_2) = 0$ и на основании теоремы Лагранжа

$$F(x, y_2) - F(x, y_1) = 0 = (y_2 - y_1)\,F'_y(x, y),$$

где $y_1 < y < y_2$, откуда $F'_y(x, y) = 0$, что по нашему предположению невозможно.

Применим теперь найденное правило дифференцирования неявных функций к простейшей задаче на так называемый «условный экстремум»; обобщение этой задачи на случай большего числа переменных составляет собой важную и интересную теорию условно экстремальных значений, которую мы не можем, к сожалению, развернуть здесь во всей широте.

Рассмотрим дифференцируемую в некоторой области функцию $F(x, y)$ и будем считать величину x независимой переменной, а величину y — функцией от x, определяемой на некотором отрезке $[a, b]$ соотношением

$$\Phi(x, y) = 0, \tag{19}$$

где функция $\Phi(x, y)$ также дифференцируема; таким образом, $F(x, y)$ фактически является функцией одной независимой переменной x, заданной указанным сложным образом; и в анализе, и в приложениях очень часто приходится исследовать функции, заданные именно таким способом. Разыскиваются экстремальные значения заданной функции внутри отрезка $[a, b]$.

Согласно общей теории эти экстремальные значения (относительные максимумы и минимумы) должны обращать в нуль производную данной функции; мы должны, таким образом, приравнять нулю производную данной функции $z = F(x, y)$ по x, считая при этом, конечно, y функцией от x, определяемой соотношением (19), т. е. приравнять нулю так называемую «полную» производную $\frac{dz}{dx}$ этой функции по x:

$$\frac{dz}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0. \tag{20}$$

В этом уравнении $\frac{\partial F}{\partial x}$ и $\frac{\partial F}{\partial y}$ являются функциями от x и y; для определения величины $\frac{dy}{dx}$ мы должны, конечно, применить правило дифференцирования неявных функций; определяющее функцию y соотношение (19)

дает

$$\frac{dy}{dx} = -\frac{\frac{\partial \Phi}{\partial x}}{\frac{\partial \Phi}{\partial y}},$$

в результате чего уравнение (20) приводится к виду

$$\frac{\partial F}{\partial x}\frac{\partial \Phi}{\partial y} - \frac{\partial \Phi}{\partial x}\frac{\partial F}{\partial y} = 0;$$

это уравнение вместе с уравнением (19) и позволяет определить все «критические» пары (x, y), т. е. все точки, в которых $\frac{dz}{dx} = 0$. Дальнейшего исследования мы здесь касаться не можем.

ЛЕКЦИЯ VI

ИНТЕГРАЛ

Введение.— Определение интеграла.— Условия интегрируемости.— Схема геометрических и физических приложений.— Связь с дифференцированием.— Теоремы о средних значениях.— Обобщенные интегралы.— Двойные интегралы.— Вычисление двойных интегралов.— Общая идея интеграла.

Введение. Интегральное исчисление первоначально развивалось независимо от учения о дифференцировании. Только к концу 17-го столетия, когда обе ветви получили уже значительное развитие и научились решать — каждая своими собственными методами — большое число задач геометрии и механики, была со всей полнотой раскрыта существовавшая между ними глубокая связь: их основные проблемы оказались двумя взаимно обратными задачами анализа бесконечно малых; интегрирование и дифференцирование функций стоят друг к другу в том же отношении, как сложение и вычитание чисел.

Этот исторический момент обычно считают датой рождения того учения, которое в наше время называют математическим анализом. И действительно, именно с этого времени, когда главным движущим рычагом как дифференциального, так и интегрального исчисления сделалась эта идея их неразрывной принципиальной связанности, развитие обеих ветвей пошло с исключительной быстротой; в особенности интегральное исчисление с этого момента получило возможность перейти от решения отдельных разрозненных задач к созданию достаточно мощных общих методов.

Эта своеобразная историческая судьба взаимосвязи между двумя основными разделами анализа до сих пор

находит отражение в системах изложения их на страницах учебных руководств: в то время как одни авторы в целях большей логической стройности определяют интеграцию функций как операцию, обратную дифференцированию, другие предпочитают, до известной степени воспроизводя историческое развитие, определять обе операции независимо друг от друга, чтобы лишь позднее установить взаимную связь их. С *формальной* стороны, конечно, безразлично, какой из этих двух путей мы выберем; но и *по существу* трудно, даже, пожалуй, невозможно утверждать, что тот или другой из этих путей заслуживает предпочтения перед другим; дело в том, что здесь многое, в сущности, почти все, зависит от запросов и интересов учащихся: тот, в ком логический, идейный интерес к математике доминирует над прикладным, практическим, будет приветствовать введение интегрирования как операции, обратной дифференцированию, ибо, может быть, сам он уже по свойственной математикам привычке, учась дифференцировать, думал о том, как выглядит обратная операция; напротив, для учащегося с прикладным направлением интересов этот путь может показаться вычурным; изучение обратной операции как таковой, когда не видно еще, для каких конкретных задач она может оказаться полезной, естественно, покажется ему недостаточно аргументированным.

Определение интеграла. Понятие интеграла возникло и постепенно укрепилось в своем значении, когда целый ряд задач геометрии и механики приводил к необходимости производить над функциями одну и ту же аналитическую операцию, содержание которой составлял предельный переход некоторого совершенно определенного типа. Природа этой операции вам, конечно, хорошо известна. Пусть задана некоторая функция $f(x)$ на отрезке $[a, b]$. Разобьем этот отрезок на n частей, обозначая точки деления через

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b;$$

в каждом из отрезков $[x_{k-1}, x_k]$ $(k = 1, 2, \ldots, n)$ выберем произвольную точку ξ_k; значение функции $f(x)$ в точке ξ_k умножим на длину $x_k - x_{k-1}$ соответствующего

отрезка и составим сумму всех таких произведений

$$\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}); \tag{1}$$

представим себе теперь, что перед нами вся совокупность подобного рода разбиений отрезка $[a, b]$ и всех возможных выборов точек ξ_k (причем и число n частичных отрезков можно как угодно варьировать), а значит, и всех возможных значений суммы (1); обозначим через l длину наибольшего из частичных отрезков $[x_{k-1}, x_k]$ в данном разбиении. Если существует такое число I, что сумма (1) стремится к I при $l \to 0$, каковы бы ни были предпринятые разбиения и выбранные точки ξ_k, то это число называют интегралом функции $f(x)$ на отрезке $[a, b]$ и обозначают через

$$\int_a^b f(x)\,dx.$$

Несколько более точная формулировка гласит: число I называется интегралом функции $f(x)$ на отрезке $[a, b]$, если, как бы мало ни было $\varepsilon > 0$, найдется такое другое положительное число δ, что при любом разбиении отрезка $[a, b]$, удовлетворяющем требованию $l < \delta$, и при любом выборе точек ξ_k

$$\left| I - \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) \right| < \varepsilon.$$

Как бы то ни было, мы видим, что здесь речь идет о своеобразном и довольно сложном предельном переходе; процесс, лежащий в основе этого предельного перехода, трудно описать (как мы это делаем обычно) указанием на поведение одной какой-либо независимой переменной; можно за такую переменную принять l (и описывать процесс символом $l \to 0$), но при этом нельзя терять из виду, что сумма (1), о пределе которой идет речь, не есть однозначная функция величины l, ибо данному значению l, очевидно, соответствует бесчисленное множество различных разбиений, а при выбранном разбиении — еще бесчисленное множество

возможностей в отношении выбора точек ξ_k; для существования предела I нужно, чтобы вся эта совокупность значений суммы (1) при достаточно малом l была как угодно близка к I.

Такая сложность в процессе, лежащем в основе интеграции, приводит к некоторым неудобствам; и в общих теоретических построениях, и в конкретно-практических случаях требуемая определением независимость предела I от характера разбиений и выбора точек ξ_k хотя и может быть строго доказана, но доказательства эти своей формальной громоздкостью часто очень обременяют рассуждения. Вот почему во многих современных изложениях принято определять интеграл несколько видоизмененным путем, избегающим на первых порах какого бы то ни было предельного перехода. К этому новому определению (которое, конечно, формально эквивалентно предыдущему) мы теперь и переходим.

Пусть $f(x)$ — любая ограниченная функция на отрезке $[a, b]$; обозначим ее грани через M (верхнюю) и m (нижнюю). Для любого разбиения T отрезка $[a, b]$ обозначим через M_k и m_k соответственно верхнюю и нижнюю грани функции $f(x)$ на отрезке $[x_{k-1}, x_k]$ и положим $x_k - x_{k-1} = \Delta_k$ $(1 \leqslant k \leqslant n)$. Пусть, далее,

$$S_T = \sum_{k=1}^{n} M_k \Delta_k,$$

$$s_T = \sum_{k=1}^{n} m_k \Delta_k;$$

таким образом, каждому разбиению T соответствует определенная «верхняя» сумма S_T и определенная «нижняя» сумма s_T; при этом, очевидно, $m \leqslant m_k \leqslant \leqslant M_k \leqslant M$ $(1 \leqslant k \leqslant n)$, так что при любом разбиении T будет $m(b-a) \leqslant s_T \leqslant S_T \leqslant M(b-a)$; совокупность значений, принимаемых как верхней суммой S_T, так и нижней суммой s_T при всевозможных разбиениях отрезка $[a, b]$, есть, таким образом, ограниченное множество.

Нижнюю грань множества всех верхних сумм S_T мы называем верхним интегралом функции $f(x)$ на отрезке

$[a, b]$ (или в пределах от a до b) и обозначаем через

$$\bar{I} = \bar{\int_a^b} f(x)\,dx.$$

Подобным же образом верхнюю грань множества всех нижних сумм s_T мы называем нижним интегралом функции $f(x)$ на отрезке $[a, b]$ (или в пределах от a до b) и обозначаем через

$$\underline{I} = \underline{\int_a^b} f(x)\,dx.$$

Таким образом, верхним и нижним интегралами на данном отрезке обладает всякая ограниченная функция; эти два интеграла определяются, как вы видите, без всякого предельного процесса, как грани некоторых множеств.

Если верхний и нижний интегралы функции $f(x)$ на отрезке $[a, b]$ совпадают между собой, то общее значение их называется интегралом данной функции в пределах от a до b и обозначается через

$$I = \int_a^b f(x)\,dx;$$

функция $f(x)$ в этом случае называется интегрируемой на отрезке $[a, b]$.

Для того чтобы сделать из этого определения нужные нам выводы, нам понадобится несколько совсем элементарных вспомогательных предложений. Условимся говорить, что разбиение T' отрезка $[a, b]$ является *продолжением* разбиения T, если все точки деления разбиения T служат вместе с тем и точками деления для разбиения T' (но, вообще говоря, T' содержит и новые точки деления, не содержащиеся в T).

Лемма I. *Если T' есть продолжение разбиения T, то*

$$S_{T'} \leqslant S_T, \quad s_{T'} \geqslant s_T.$$

В самом деле, при переходе от разбиения T к его продолжению T' каждый член $M_k\Delta_k$ суммы S_T

заменяется группой членов

$$\sum_{r=1}^{s} M_{k_r}\Delta_{k_r},$$

где $\Delta_{k_1}, \Delta_{k_2}, \ldots, \Delta_{k_r}$ — длины тех отрезков, на которые разбивается отрезок $[x_{k-1}, x_k]$ при переходе от разбиения T к разбиению T', а M_{k_r} — верхняя грань функции $f(x)$ на отрезке Δ_{k_r}; так как, очевидно, $M_{k_r} \leqslant M_k$ $(1 \leqslant r \leqslant s)$, то

$$\sum_{r=1}^{s} M_{k_r}\Delta_{k_r} \leqslant M_k \sum_{r=1}^{s} \Delta_{k_r} = M_k\Delta_k,$$

т. е. та группа членов, которой заменяется $M_k\Delta_k$, имеет сумму, не бо́льшую чем $M_k\Delta_k$, а так как это справедливо для всех k, то $S_{T'} \leqslant S_T$; совершенно аналогично доказывается и второе из утверждаемых неравенств.

Лемма II. *Для любых разбиений T_1 и T_2 отрезка* $[a, b]$

$$S_{T_1} \geqslant s_{T_2}.$$

В самом деле, совокупность всех точек деления обоих разбиений T_1 и T_2 определяет некоторое третье разбиение T, являющееся, очевидно, продолжением каждого из разбиений T_1 и T_2; но для одного и того же разбиения T, очевидно, $S_T \geqslant s_T$; поэтому в силу леммы I

$$S_{T_1} \geqslant S_T \geqslant s_T \geqslant s_{T_2},$$

что и требовалось доказать.

Непосредственным следствием леммы II является

Лемма III. *Для любой (ограниченной) функции* $\overline{I} \geqslant \underline{I}$.

В самом деле, так как любая из сумм s_T не превосходит любой из сумм S_T, то и верхняя грань $\underline{I}$ совокупности всех s_T не может превосходить нижней грани $\overline{I}$ совокупности всех S_T.

Пусть теперь l_T означает длину наибольшего из частичных отрезков, на которые распадается отрезок $[a, b]$ в силу разбиения T.

Лемма IV. *При любом $\varepsilon > 0$ существует такое $\lambda > 0$, что для всех разбиений, для которых $l_T < \lambda$, мы имеем $S_T < \overline{I} + \varepsilon$, $s_T > \underline{I} - \varepsilon$.*

Это предложение иногда выражают несколько короче, говоря, что $S_T \to \overline{I}$ и $s_T \to \underline{I}$, когда $l_T \to 0$; такая редакция не может привести к смешениям, если только помнить, что S_T и s_T не являются однозначными функциями от l_T.

Доказательство. Не ограничивая общности рассуждения, мы можем допустить, что $f(x) \geqslant 0$ при $a \leqslant x \leqslant b$ (этого мы можем всегда достигнуть прибавлением к $f(x)$ достаточно большого числа A, причем все интегралы и суммы увеличиваются на $A(b-a)$). По определению нижней грани существует разбиение T_0, для которого $S_{T_0} < \overline{I} + \frac{\varepsilon}{2}$. Обозначим через $x_1, x_2, \ldots, x_n$ внутренние точки деления отрезка $[a, b]$ при этом разбиении, и пусть M — верхняя грань функции $f(x)$ на отрезке $[a, b]$. Положим, наконец, $\lambda = \frac{\varepsilon}{4nM}$, и пусть T — любое разбиение отрезка $[a, b]$, для которого $l_T < \lambda$. Мы покажем, что $S_T < \overline{I} + \varepsilon$.

С этой целью разобьем частичные отрезки, на которые распадается отрезок $[a, b]$ при разбиении T, на две группы; к первой группе отнесем всякий отрезок, целиком принадлежащий одному из отрезков $[x_k - \lambda, x_k + \lambda]$ $(1 \leqslant k \leqslant n)$, а ко второй группе — остальные частичные отрезки (рис. 25). Соответственно этому и сумма

Рис. 25.

S_T распадается на суммы S_T^{I} и S_T^{II}. В сумме S_T^{I} все члены имеют первые множители, не превосходящие M, в то время как сумма длин отрезков, составляющих вторые множители, не более чем $2n\lambda = \frac{\varepsilon}{2M}$, так что

$$S_T^{\mathrm{I}} \leqslant M \cdot \frac{\varepsilon}{2M} = \frac{\varepsilon}{2}.$$

С другой стороны, часть суммы S_T^{II}, относящаяся к отрезкам разбиения T, содержащимся внутри отрезка $[x_{k-1}, x_k]$ (рис. 25), имеет первые множители, не большие чем M_k, а сумму вторых — не бóльшую чем $x_k - x_{k-1} = \Delta_k$, вследствие чего величина этой части не больше чем $M_k\Delta_k$, а значит, $S_T^{II} \leqslant \sum_{k=1}^{n} M_k\Delta_k = S_{T_0}$.

Таким образом,

$$S_T = S_T^{I} + S_T^{II} \leqslant S_T^{I} + S_{T_0} < \frac{\varepsilon}{2} + \bar{I} + \frac{\varepsilon}{2} = \bar{I} + \varepsilon,$$

что мы и хотели доказать.

Совершенно аналогично доказывается, конечно, и второе из утверждаемых неравенств.

Теорема. *Пусть функция $f(x)$ интегрируема на отрезке $[a, b]$; тогда, сколь бы мало ни было $\varepsilon > 0$, существует такое положительное число λ, что при любом разбиении T, для которого $l_T \leqslant \lambda$, мы будем иметь, как бы мы ни выбирали точки $\xi_k \in [x_{k-1}, x_k]$, $k = 1, 2, \ldots, n$,*

$$\left| \sum_{k=1}^{n} f(\xi_k)\Delta_k - I \right| < \varepsilon,$$

где $I = \int_a^b f(x)\,dx$.

В самом деле, так как, очевидно,

$$m_k \leqslant f(\xi_k) \leqslant M_k \qquad (1 \leqslant k \leqslant n)$$

при любом выборе точек ξ_k в соответствующих частичных отрезках, то

$$s_T \leqslant \sum_{k=1}^{n} f(\xi_k)\Delta_k \leqslant S_T;$$

но, с другой стороны, в силу леммы IV для всех разбиений с достаточно малым l_T

$$I - \varepsilon = \underline{I} - \varepsilon < s_T \leqslant S_T < \bar{I} + \varepsilon = I + \varepsilon;$$

сопоставляя эти неравенства с предыдущими, находим,

что для всех разбиений с достаточно малым l_T

$$I - \varepsilon < \sum_{k=1}^{n} f(\xi_k)\,\Delta_k < I + \varepsilon,$$

что и требовалось доказать.

Обратная теорема, как легко видеть, также верна; в самом деле, если при любом разбиении с достаточно малым l_T и при любом выборе точек ξ_k сумма $\sum\limits_{k=1}^{n} f(\xi_k)\,\Delta_k$ как угодно близка к числу I, то и S_T при достаточно малом l_T будет меньше, чем $I + \varepsilon$; с другой стороны, всегда $S_T \geqslant \bar{I}$, откуда $\bar{I} < I + \varepsilon$, а значит, $\bar{I} \leqslant I$, так как ε произвольно мало; подобным же образом $\underline{I} \geqslant I$, так что $\bar{I} \leqslant \underline{I}$, откуда, в силу леммы III, $\bar{I} = \underline{I} = I$, т. е. функция $f(x)$ интегрируема на отрезке $[a, b]$.

Таким образом, данное нами новое определение интеграла, действительно, в точности равносильно первоначальному.

Примечание. Из леммы IV, очевидно, вытекает: если мы обозначим через $\omega_k = M_k - m_k$ колебание функции $f(x)$ на отрезке $[x_{k-1}, x_k]$, то *для того, чтобы функция* $f(x)$ *была интегрируема на отрезке* $[a, b]$, *необходимо и достаточно, чтобы сумма*

$$\sum_{k=1}^{n} \omega_k \Delta_k = S_T - s_T \qquad (2)$$

была как угодно мала для всех разбиений T *с достаточно малым* l_T.

Условия интегрируемости. Сам по себе указанный в последнем примечании необходимый и достаточный признак интегрируемости еще мало удобен, так как в большинстве случаев для конкретно заданных функций нелегко бывает непосредственно судить о поведении суммы (2). Однако с помощью этого признака очень легко устанавливать общие условия интегрируемости для более или менее широких классов функций.

Мы теперь и пойдем по этому пути.

Покажем прежде всего, что *всякая функция* $f(x)$, *непрерывная на отрезке* $[a, b]$, *интегрируема на этом*

отрезке. В самом деле, по свойству равномерной непрерывности (теорема 4 лекции III, стр 79) для каждого положительного числа ε найдется такое положительное число λ, что колебание функции $f(x)$ будет меньше ε на любом отрезке, длина которого меньше λ. Но тогда при любом разбиении T, для которого $l_T < \lambda$, все ω_k в сумме (2) будут меньше, чем ε, вследствие чего

$$\sum_{k=1}^{n} \omega_k \Delta_k < \varepsilon \sum_{k=1}^{n} \Delta_k = \varepsilon (b - a),$$

т. е. сумма (2) становится как угодно малой при достаточно малом l_T, откуда, как мы знаем, и вытекает интегрируемость функции $f(x)$ на отрезке $[a, b]$.

Может ли разрывная функция оказаться интегрируемой? Легко видеть на примерах, что может; легко даже установить в этом направлении и некоторые общие закономерности. Покажем, например, что *ограниченная функция* $f(x)$, *имеющая на отрезке* $[a, b]$ *только одну точку разрыва* c, *интегрируема на этом отрезке* (каков бы ни был характер разрыва в точке c).

С этой целью обозначим через μ верхнюю грань функции $|f(x)|$ на отрезке $[a, b]$ и выберем произвольное положительное число ε; на отрезках $\left[a, c - \frac{\varepsilon}{2\mu}\right]$ и $\left[c + \frac{\varepsilon}{2\mu}, b\right]$ функция $f(x)$ непрерывна; мы можем поэтому снова найти такое положительное число λ, что колебание функции на любом отрезке длины $<\lambda$, целиком помещающемся либо в отрезке $\left[a, c - \frac{\varepsilon}{2\mu}\right]$, либо в отрезке $\left[c + \frac{\varepsilon}{2\mu}, b\right]$, будет меньше чем ε. Пусть теперь снова T есть любое разбиение отрезка $[a, b]$, для которого $l_T < \lambda$; среди отрезков Δ_k будут, вообще говоря, такие, которые целиком помещаются в одном из двух вышеупомянутых отрезков, а также и такие, которые хотя бы частично захватывают отрезок $\left[c - \frac{\varepsilon}{2\mu}, c + \frac{\varepsilon}{2\mu}\right]$ (рис. 26). Для отрезков Δ_k первого типа колебания ω_k в сумме (2) меньше чем ε, вследствие чего относящаяся

к отрезкам первого типа часть суммы (2) меньше чем $\varepsilon(b-a)$. Для отрезков Δ_k второго типа относительно колебаний ω_k мы можем только утверждать, что каждое из них не больше чем 2μ; но так как все отрезки второго типа целиком принадлежат отрезку $\left[c-\frac{\varepsilon}{2\mu}-\lambda, c+\frac{\varepsilon}{2\mu}+\lambda\right]$, то сумма длин их не превосходит $\frac{\varepsilon}{\mu}+2\lambda$, а следовательно, соответствующая им часть суммы (2) не больше чем $2\mu\left(\frac{\varepsilon}{\mu}+2\lambda\right)=2\varepsilon+4\lambda\mu$; сопоставляя это

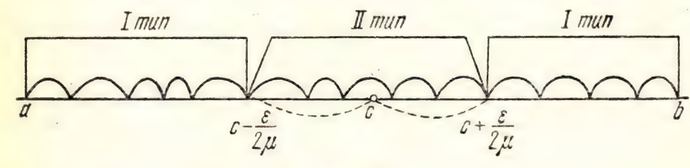


Рис. 26.

с найденной выше оценкой для первой части суммы (2), мы находим, что при $l_T<\lambda$

$$\sum_{k=1}^{n}\omega_k\Delta_k<\varepsilon(b-a)+2\varepsilon+4\lambda\mu;$$

так как ε и λ сколь угодно малы, то этим установлена интегрируемость функции $f(x)$ на отрезке $[a, b]$.

Мы сознательно провели с такой подробностью это несложное рассуждение; здесь важно уловить основную мысль доказательства; сумма (2) оказывается малой потому, что: 1) в областях, где функция непрерывна, малы *первые* множители ω_k (равномерная непрерывность!) и 2) отрезки, заключающие точки разрыва, дают малую сумму *вторых* множителей; отсюда малыми оказываются обе части суммы (2), а значит, и вся эта сумма. Теперь легко понять (и строго доказать, конечно), что и любое конечное число разрывов функции $f(x)$ на отрезке $[a, b]$ не помешает ее интегрируемости, если только она остается ограниченной. Напротив, если точки разрыва заполняют слишком большую долю отрезка $[a, b]$, то функция может стать неинтегрируе-

мой. Так, функция Дирихле (лекция III, стр 65), имеющая разрыв в каждой точке, не интегрируема ни на каком отрезке; в самом деле, на любом отрезке Δ_k для этой функции $\omega_k = 1$, вследствие чего при любом разбиении отрезка $[a, b]$

$$\sum_{k=1}^{n} \omega_k \Delta_k = \sum_{k=1}^{n} \Delta_k = b - a.$$

Занимающий нас круг вопросов находит полное разрешение всех относящихся сюда задач в следующем предложении.

Теорема. *Для того чтобы ограниченная функция $f(x)$ была интегрируема на отрезке $[a, b]$, необходимо и достаточно выполнение следующего условия: как бы мало ни было $\varepsilon > 0$, все точки отрезка $[a, b]$, в которых колебание функции $f(x)$ превосходит ε, могут быть заключены в конечное число отрезков, сумма длин которых меньше чем ε.*

Доказательство. 1) Пусть условие теоремы выполнено. Пусть $\delta_1, \delta_2, \ldots, \delta_n$ — система отрезков, заключающих внутри себя все точки, в которых колебание функции $f(x)$ превосходит ε; пусть при этом $\sum_{i=1}^{n} \delta_i < \varepsilon$; систему дополнительных отрезков будем обозначать через $d_1, d_2, \ldots, d_N$. Так как в любой точке любого отрезка d_i колебание функции $f(x)$ не превосходит ε, то, согласно теореме на стр. 86 лекции III, колебание этой функции будет меньше чем 2ε на любом отрезке достаточно малой длины, целиком лежащем в одном из отрезков d_i. Теперь остается только вообразить себе любое достаточно мелкое (т. е. обладающее достаточно малым l_T) разбиение отрезка $[a, b]$ на частичные отрезки Δ_k и провести дальнейшее рассуждение по схеме стр. 177—178; разбивая сумму (2) на две части соответственно тому, принадлежит ли отрезок Δ_k целиком одному из отрезков d_i или хотя бы частично захватывает один из отрезков δ_i, мы легко находим, что первая часть суммы (2) меньше чем $2\varepsilon(b - a)$, а вторая не превосходит $2\mu(\varepsilon + 2nl_T)$, где μ — верхняя грань функции $|f(x)|$ на отрезке $[a, b]$, а n — число отрезков δ_i;

выбирая ε и l_T достаточно малыми, мы можем, таким образом, сделать сумму (2) как угодно малой, что и требовалось доказать.

2) Пусть функция $f(x)$ интегрируема на отрезке $[a, b]$, и пусть ε — любое положительное число; мы можем выбрать такое разбиение отрезка $[a, b]$, чтобы соответствующая сумма (2) была меньше чем ε^2; заметим теперь, что все слагаемые суммы (2) неотрицательны и что, следовательно, мы не увеличим эту сумму, если сохраним в ней только те члены, в которых $\omega_k > \varepsilon$, выбросив все остальные; таким образом,

$$\varepsilon^2 > \sum_{k=1}^{n} \omega_k \Delta_k \geqslant \sum_{\omega_k > \varepsilon} \omega_k \Delta_k \geqslant \varepsilon \sum_{\omega_k > \varepsilon} \Delta_k,$$

откуда

$$\sum_{\omega_k > \varepsilon} \Delta_k < \varepsilon;$$

но отрезки Δ_k, входящие в состав этой последней суммы, очевидно, содержат все, кроме, быть может, конечного числа, точки отрезка $[a, b]$, в которых колебание функции $f(x)$ превосходит ε (ибо такая точка не может быть внутренней точкой отрезка Δ_k, для которого $\omega_k \leqslant \varepsilon$); таким образом, условие интегрируемости, сформулированное в доказываемой нами теореме, действительно является необходимым.

Из доказанной теоремы, в частности, следует *интегрируемость всякой монотонной функции* на отрезке $[a, b]$; в самом деле, как мы видели в лекции III, всякая такая функция необходимо ограничена на данном отрезке; далее, мы видели, что, как бы мало ни было $\varepsilon > 0$, число точек отрезка $[a, b]$, в которых такая функция имеет колебание $> \varepsilon$, обязательно конечно; но конечную группу точек всегда можно, разумеется, заключить в конечную группу интервалов, сумма длин которых как угодно мала; таким образом, признак интегрируемости, о котором идет речь в только что доказанной нами теореме, необходимо выполняется для каждой монотонной функции.

Схема геометрических и физических приложений. Рассмотренное нами понятие интеграла, тесно связанное (при посредстве идеи предела или грани) с суммами определенного вида и не стоящее само по себе ни в какой связи с понятиями дифференциального исчисления, имеет, как вы знаете, большое число геометрических и физических применений. Нам нет, конечно, надобности останавливаться здесь на этих приложениях ради них самих или хотя бы составлять примерный перечень наиболее важных из них. Однако очень важно для нас проследить, что во всех случаях подобного рода применений мы встречаемся со своеобразной и несколько тонкой логической ситуацией, очень важная в принципиальном отношении специфика которой обычно недостаточно подчеркивается в руководствах.

Пусть, например, речь идет о вычислении площади «криволинейной трапеции» рис. 27, ограниченной сверху графиком функции $y = f(x)$, которую мы для простоты предположили положительной и непрерывной. Конечно, всем вам хорошо знакома эта задача, так же как и ее решение с помощью интегрального исчисления. Вдумаемся, однако, несколько подробнее в ту логическую ситуацию, в которой мы здесь находимся. Вычислять площадь той или другой фигуры имеет, очевидно, смысл лишь тогда, если самое понятие площади точно определено. В данный момент, когда мы ищем метод, который позволил бы вычислить площадь криволинейной трапеции рис. 27, владеем ли мы уже таким определением, знаем ли мы точное определение той величины, которую мы собираемся вычислить? Конечно, нет; мы знаем определение площади лишь для прямолинейных фигур (многоугольников) и некоторых частей круга; кривая же $y = f(x)$, ограничивающая нашу трапецию, вообще говоря, ничего общего не имеет с окружностью.

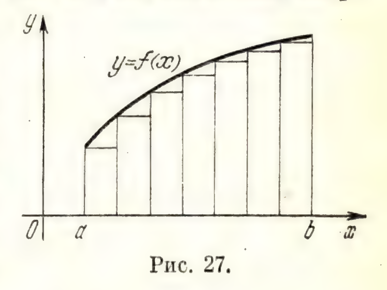


Рис. 27.

Как же мы можем, не зная, что такое искомая площадь, приступить к ее вычислению? И — что всего удивительнее — как же это вычисление нам все-таки удается, хотя мы и не знаем даже, что именно мы вычисляем? Дело заключается в том, что фактически мы ставим и решаем здесь задачу гораздо более ответственную, чем простое вычисление; мы одновременно *определяем* понятие площади для нашей криволинейной трапеции и находим метод вычисления этой площади. Когда мы утверждаем, что площадь нашей фигуры равна пределу площадей тех прямолинейных фигур лестничного вида, одна из которых изображена на рис. 27, то такое утверждение является не теоремой, а *определением* площади нашей криволинейной трапеции; пытаться доказывать это утверждение было бы поэтому бессмысленным; напротив, утверждение, что упомянутый предел *существует* при тех или других условиях (например, в случае непрерывности функции $f(x)$), есть *теорема*, которая может и должна быть доказана. И в точности такой же характер имеет логическая природа любой из подобного рода геометрических и физических задач; собираемся ли мы вычислять длину криволинейной дуги, объем или поверхность тела вращения, или работу данной силы на данном участке пути и т. д.— во всех случаях речь идет об отыскании количественной меры для такого понятия, которое до этого момента оказывается определенным лишь в некоторых простейших частных случаях; задача состоит в том, чтобы дать этому понятию целесообразное общее определение, которое одновременно содержало бы метод получения количественной оценки для соответствующей данному понятию величины.

Спросим себя теперь, какие же общие черты перечисленных выше (и, конечно, многих других) задач геометрии и физики обусловливают возможность аналитически одинаково решать эти задачи, используя во всех случаях именно интеграл в качестве орудия, логически определяющего искомую величину и вместе с тем дающего ей количественную оценку. Таких основных черт можно отметить две. Во всех случаях искомая величина зависит от некоторого отрезка $[a, b]$, на кото-

рый она «распространяется» и с изменением которого она меняется. Так, на рис. 27 мы получим другую площадь, если отрезок $[a, b]$ заменим другим отрезком (сохраняя, конечно, функцию $f(x)$ неизменной); работа силы различна на различных участках пути движущейся точки и т. д. С другой стороны, изучаемая величина в каждой конкретной задаче зависит от некоторой функции $f(x)$; в случае рис. 27 — это ордината той точки верхней стороны рассматриваемой трапеции, абсцисса которой равна x; в случае вычисления работы — это величина силы, действующей на расстоянии x от начала отсчета, и т. д. Таким образом, для того чтобы задача рассматриваемого нами типа получила определенную постановку, необходимо прежде всего задание некоторой функции $f(x)$ и некоторого отрезка $a \leqslant x \leqslant b$, к которому мы относим нашу задачу; можно сказать, что та величина, определение и численную оценку которой мы ищем, есть функция $V(f; a, b)$ трех элементов, которые могут быть выбираемы независимо друг от друга: функции $f(x)$ и чисел a и b. Применение интеграла как метода решения всех перечисленных (и многих других) задач в основном обусловлено следующими свойствами этой зависимости $V(f; a, b)$.

1) Как функция отрезка $[a, b]$ величина V *аддитивна*, т. е. при $a < c < b$

$$V(f; a, b) = V(f; a, c) + V(f; c, b).$$

Действительно, площадь фигуры рис. 27 при всяком разумном определении, очевидно, должна равняться сумме площадей криволинейных трапеций, на которые она распадается при разбиении отрезка $[a, b]$ на частичные отрезки. Объем тела вращения равен сумме объемов, расположенных вокруг отдельных участков, на которые мы разбиваем ось вращения; длина дуги кривой равна сумме длин ее участков; работа силы на данном пути равна сумме работ, производимых ею на различных участках этого пути при любом его разбиении, и т. д.

2) Если функция $f(x)$ на участке $[a, b]$ постоянна, $f(x) = C$, то

$$V(f; a, b) = C(b - a).$$

Действительно, если $f(x) = C$ при $a \leqslant x \leqslant b$, то фигура рис. 27 есть прямоугольник, площадь которого равна $C(b-a)$; если в случае тела вращения $f(x)$ означает площадь поперечного сечения и если $f(x) = C$ при $a \leqslant x \leqslant b$, то мы имеем дело с цилиндром, объем которого равен $C(b-a)$; если сила, которая действует на точку, сохраняет на пути $a \leqslant x \leqslant b$ постоянное значение C, то работа этой силы на данном пути будет $C(b-a)$, и т. д.

Легко теперь убедиться, что во всех случаях, когда зависимость искомой величины V от данных элементов $f(x)$, a и b обладает чертами 1) и 2), естественно ожидать, что задача будет иметь своим решением интеграл

$$V = \int_a^b f(x)\,dx.$$

В самом деле, если мы данный отрезок $[a, b]$ разобьем, как обычно, на части с помощью точек деления

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b,$$

то прежде всего в силу свойства 1)

$$V(f;\, a, b) = \sum_{k=1}^{n} V(f;\, x_{k-1}, x_k);$$

если бы функция $f(x)$ для всех точек x отрезка $[x_{k-1}, x_k]$ имела некоторое постоянное значение C_k, то в силу свойства 2) мы имели бы

$$V(f;\, x_{k-1}, x_k) = C_k(x_k - x_{k-1});$$

в действительности функция $f(x)$, вообще говоря, не постоянна на отрезке $[x_{k-1}, x_k]$. Но если эта функция непрерывна и если отрезок $[x_{k-1}, x_k]$ очень мал, то все значения функции $f(x)$ на этом отрезке мало отличаются друг от друга; взяв какое-нибудь одно $f(\xi_k)$ из этих значений, мы можем сказать, что на всем отрезке $[x_{k-1}, x_k]$ функция $f(x)$ приблизительно равна $f(\xi_k)$, вследствие чего величину $V(f;\, x_{k-1}, x_k)$ нам естественно принять (не надо забывать, что эта величина еще не определена) приблизительно равной $f(\xi_k)\ (x_k - x_{k-1})$,

так что приближенно

$$V(f;\, a,\, b) \approx \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}),$$

где ξ_k — любая точка отрезка $[x_{k-1},\, x_k]$. При этом мы, естественно, полагаем, что погрешность этого приближенного равенства тем меньше, чем меньше отрезки $[x_{k-1},\, x_k]$, т. е. чем меньше l_T для предпринятого разбиения; но отсюда уже совершенно естественно определять точное значение величины V как

$$\lim_{l_T \to 0} \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) = \int_a^b f(x)\, dx.$$

Связь с дифференцированием. На первом этапе развития интегрального исчисления, когда огромное значение связи между интегрированием и дифференцированием не вошло еще в должной мере в идейный мир ученых, задачи описанного нами типа решались прямым вычислением интеграла как предела суммы.

Выбором особенно удобных разбиений и специально подобранных значений ξ_k старались в каждом отдельном случае, т. е. применительно к каждой индивидуальной функции $f(x)$, по возможности облегчить и упростить это, вообще говоря, очень громоздкое вычисление.

Некоторые задачи были решены таким путем уже в древности, позднее к ним присоединился ряд новых успехов; и, однако, все же до того момента, когда основным методом вычисления интегралов сделалось использование связи между интегрированием и дифференцированием, все результаты оставались разрозненными и каждая вновь возникающая задача требовала, в сущности, создания нового метода для своего решения. Свои общие и полноценные методы интегральное исчисление приобрело только в тесном взаимодействии с учением о дифференцировании.

Вы хорошо знаете, конечно, в чем состоит связь интегрирования с дифференцированием: интеграл

$$F(x) = \int_a^x f(u)\, du$$

имеет своей производной «подынтегральную функцию» $f(x)$ в каждой точке, где эта функция непрерывна (в частности, для непрерывной функции — всюду). Доказательство очень просто: если $|h|$ достаточно мало, то при $x - |h| \leqslant u \leqslant x + |h|$ мы имеем

$$f(x) - \varepsilon \leqslant f(u) \leqslant f(x) + \varepsilon,$$

где ε — заранее выбранное, как угодно малое положительное число; отсюда

$$hf(x) - \varepsilon |h| \leqslant \int_x^{x+h} f(u)\, du \leqslant hf(x) + \varepsilon |h|,$$

и, значит,

$$f(x) - \varepsilon \leqslant \frac{1}{h} \int_x^{x+h} f(u)\, du = \frac{F(x+h) - F(x)}{h} \leqslant f(x) + \varepsilon.$$

Ввиду произвольной малости ε отсюда в самом деле вытекает

$$F'(x) = f(x).$$

В силу этой связи всякое правило дифференцирования дает некоторое правило интегрирования, будучи прочитанным, так сказать, справа налево. Но более того: многие общие приемы дифференцирования допускают по меньшей мере частичное обращение и приводят, таким образом, к важнейшим общим приемам интегрирования; так, правило дифференцирования алгебраической суммы полностью обращается и дает правило интегрирования алгебраической суммы (которое, впрочем, легко может быть выведено и из первоначального определения интеграла); правило дифференцирования произведения приводит к так называемой «интеграции по частям», силу которой вы хорошо знаете; обращением правила дифференцирования «функции от функции» служит известный вам еще более мощный прием «интеграции посредством подстановки».

Применение всей совокупности этих приемов приводит к возможности интегрировать большое число элементарных функций (в частности, *все* рациональные

функции). И если тем не менее для очень многих элементарных функций мы все же не умеем написать их интегралов, то происходит это не в силу недостаточной мощности применяемых нами методов, а по другой причине, имеющей несравненно большее принципиальное значение; если при дифференцировании элементарных функций мы во всех случаях снова приходим к элементарным функциям, то интегрирование элементарных функций представляет в этом отношении совсем другую картину; очень часто бывает, что интеграл такой функции хотя, конечно, и существует, но не является элементарной функцией и не может поэтому быть выражен никакой элементарной формулой; так обстоит дело, например, с интегралами таких простых функций, как $\frac{1}{\lg x}$ и $\frac{1}{\sqrt{x^3+1}}$. Такую новую функцию приходится исследовать, не имея для этого (по крайней мере на первых порах) никакого орудия, кроме определяющего ее интеграла.

Теоремы о средних значениях. В дифференциальном исчислении теоремой о среднем значении обычно называют теорему Лагранжа: если функция $f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема внутри него, то существует такая внутренняя точка c этого отрезка, что

$$f(b) - f(a) = f'(c)(b - a).$$

Вообще для «теорем о средних значениях» характерно это присутствие в их формулировках некоторого числа c («среднего значения» величины x между a и b), о котором известно только, что оно лежит внутри отрезка $[a, b]$, а более точно оно не характеризуется. В этом смысле и формула Коши (лекция V), и написанная с тем или другим видом остаточного члена формула Тэйлора представляют собой теоремы о средних значениях; очень часто формулируют теоремы этого рода в несколько иных обозначениях: говорят об отрезке $[a, a + h]$, а неопределенно расположенную внутреннюю точку этого отрезка обозначают через $a + \theta h$, где двойное неравенство $0 < \theta < 1$ есть все, что утверждают теоремы о числе θ; присутствие в формулировке

такого неопределенного числа, подчиненного неравенствам $0 < \theta < 1$, также может служить типичным симптомом теорем о средних значениях.

В интегральном исчислении роль теорем о средних значениях во всяком случае не меньше, чем в дифференциальном; с одной из них, так называемой *первой* теоремой о среднем значении, вы, несомненно, знакомы; в простейшем случае она утверждает, что если функция $f(x)$ непрерывна на отрезке $[a, b]$, то

$$\int_a^b f(x)\,dx = f(c)(b-a), \tag{3}$$

где c — некоторая внутренняя точка отрезка $[a, b]$; это утверждение доказывается непосредственным применением теоремы Лагранжа к функции $F(x) = \int_a^x f(u)\,du$ на отрезке $[a, b]$. Более общими являются соотношения

$$m(b-a) \leqslant \int_a^b f(x)\,dx \leqslant M(b-a) \qquad (a < b), \tag{4}$$

где m и M соответственно означают нижнюю и верхнюю грани функции $f(x)$ на отрезке $[a, b]$; эти соотношения имеют место для любой ограниченной интегрируемой функции $f(x)$; однако и в случае, когда эта функция непрерывна, равенство (3) ввиду неизвестного положения точки c дает в смысле оценки входящего в него интеграла, в сущности, не больше, чем неравенства (4).

Более общая формулировка «первой» теоремы о среднем значении гласит: если функция $\varphi(x)$ на отрезке $[a, b]$ интегрируема и нигде не отрицательна, а функция $f(x)$ непрерывна, то

$$\int_a^b f(x)\varphi(x)\,dx = f(c)\int_a^b \varphi(x)\,dx, \tag{5}$$

где c снова означает некоторую (точнее не определяемую) внутреннюю точку отрезка $[a, b]$. Для простоты будем считать функции $f(x)$ и $\varphi(x)$ непрерывными и допустим, что $\varphi(x) > 0$ при $a \leqslant x \leqslant b$; для доказа-

тельства формулы (5) (в частном случае $\varphi(x) \equiv 1$ переходящей в формулу (3)) достаточно применить формулу Коши на отрезке $[a, b]$ к функциям

$$F(x) = \int_a^x f(u)\,\varphi(u)\,du, \quad \Phi(x) = \int_a^x \varphi(u)\,du,$$

что дает

$$\frac{\int_a^b f(x)\,\varphi(x)\,dx}{\int_a^b \varphi(x)\,dx} = \frac{F(b) - F(a)}{\Phi(b) - \Phi(a)} = \frac{F'(c)}{\Phi'(c)} = \frac{f(c)\,\varphi(c)}{\varphi(c)} = f(c),$$

т. е. формулу (5).

Значительно более тонким аналитическим инструментом является, однако, так называемая *вторая* теорема о среднем значении к рассмотрению которой мы теперь переходим. Эта теорема также имеет своим объектом интеграл

$$\int_a^b f(x)\,\varphi(x)\,dx, \tag{6}$$

но только в случае, когда один из двух множителей подынтегральной функции есть функция, монотонная на отрезке $[a, b]$. Пусть функция $\varphi(x)$ — неотрицательная и невозрастающая для $a < x < b$. Произведем, как обычно, разбиение T отрезка $[a, b]$ и сохраним наши обычные обозначения. Покажем, что интеграл (6) есть предел суммы

$$\sum_{k=1}^{n} \varphi(\xi_k) \int_{x_{k-1}}^{x_k} f(x)\,dx \tag{7}$$

при $l_T \to 0$. В самом деле, обозначая через μ верхнюю грань функции $|f(x)|$ на отрезке $[a, b]$ и через Δ разность между суммой (7) и интегралом (6), мы ввиду предположенного монотонного невозрастания функции

$\varphi(x)$ будем иметь

$$|\Delta| = \left| \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} [\varphi(\xi_k) - \varphi(x)] f(x)\,dx \right| \leqslant$$

$$\leqslant \sum_{k=1}^{n} [\varphi(x_{k-1}) - \varphi(x_k)]\, \mu (x_k - x_{k-1}) \leqslant$$

$$\leqslant \mu l_T [\varphi(b) - \varphi(a)]$$

и, следовательно,

$$\Delta \to 0 \quad \text{при} \quad l_T \to 0.$$

Заметив это, преобразуем теперь сумму (7) с помощью леммы Абеля (лекция IV, стр. 124); полагая

$$\int_a^{x_k} f(x)\,dx = A_k \qquad (k = 0, 1, 2, \ldots, n),$$

мы находим

$$S = \sum_{k=1}^{n} \varphi(\xi_k) \int_{x_{k-1}}^{x_k} f(x)\,dx = \sum_{k=1}^{n} \varphi(\xi_k)(A_k - A_{k-1}) =$$

$$= \sum_{k=1}^{n-1} A_k [\varphi(\xi_k) - \varphi(\xi_{k+1})] + A_n \varphi(\xi_n); \quad (8)$$

обозначим теперь через M и m грани функции

$$\int_a^x f(u)\,du \quad (9)$$

на отрезке $a \leqslant x \leqslant b$; очевидно, что $m \leqslant A_k \leqslant M$ при $0 \leqslant k \leqslant n$, а так как в силу предположенных свойств функции $\varphi(x)$ все A_k в правой части формулы (8) умножаются на неотрицательные множители, то эта формула дает

$$m\varphi(\xi_1) \leqslant S \leqslant M\varphi(\xi_1);$$

но при $l_T \to 0$ мы имеем

$$\varphi(\xi_1) \to \varphi(a+0), \quad S \to \int_a^b f(x)\,\varphi(x)\,dx,$$

вследствие чего

$$m\varphi(a+0) \leqslant \int_a^b f(x)\,\varphi(x)\,dx \leqslant M\varphi(a+0),$$

или, если $\varphi(a+0) \neq 0$ (случай $\varphi(a+0) = 0$, разумеется, тривиален),

$$m \leqslant \frac{1}{\varphi(a+0)} \int_a^b f(x)\,\varphi(x)\,dx \leqslant M;$$

но функция (9), будучи непрерывной, должна в некоторой внутренней точке ξ отрезка $[a, b]$ принимать промежуточное между ее гранями значение, стоящее в средней части этих неравенств; отсюда

$$\int_a^b f(x)\,\varphi(x)\,dx = \varphi(a+0) \int_a^\xi f(x)\,dx \qquad (a < \xi < b). \quad (10)$$

Эта формула и составляет собой содержание второй теоремы о среднем значении в наших специальных условиях (т. е. когда функция $\varphi(x)$ — неотрицательная и невозрастающая, а $f(x)$ — любая интегрируемая). «Среднее значение» ξ выступает здесь в своеобразной роли одного из пределов интеграции.

Если предположить функцию $\varphi(x)$ снова неотрицательной, но на этот раз неубывающей, то, полагая $x = b - y$, $\varphi(b-y) = \psi(y)$, мы найдем

$$\int_a^b f(x)\,\varphi(x)\,dx = \int_0^{b-a} f(b-y)\,\psi(y)\,dy;$$

так как функция $\psi(y)$, очевидно, неотрицательная и невозрастающая при $0 < y < b - a$, то мы можем применить к последнему интегралу формулу (10):

$$\int_0^{b-a} f(b-y)\,\psi(y)\,dy = \psi(+0) \int_0^\eta f(b-y)\,dy =$$

$$= \varphi(b-0) \int_{-\eta}^b f(x)\,dx,$$

где $0 < \eta < b - a$ и, следовательно, $a < \xi = b - \eta < b$;

таким образом, в этом случае вторая теорема о среднем значении принимает вид

$$\int_a^b f(x)\,\varphi(x)\,dx = \varphi(b-0)\int_\xi^b f(x)\,dx.$$

Будем, наконец, предполагать функцию $\varphi(x)$ монотонной, например невозрастающей, но о знаке ее не будем делать никаких предположений; так как функция $\varphi(x) - \varphi(b-0)$ на отрезке $[a, b]$ неотрицательная и невозрастающая, то, применяя формулу (10), находим

$$\int_a^b f(x)\,[\varphi(x) - \varphi(b-0)]\,dx = \\ = [\varphi(a+0) - \varphi(b-0)]\int_a^\xi f(x)\,dx,$$

откуда

$$\int_a^b f(x)\,\varphi(x)\,dx = \varphi(a+0)\int_a^\xi f(x)\,dx + \\ + \varphi(b-0)\int_\xi^b f(x)\,dx.$$

Таково выражение второй теоремы о среднем значении в общем случае; в частности, если функция $\varphi(x)$ непрерывна в точках a и b,

$$\int_a^b f(x)\,\varphi(x)\,dx = \varphi(a)\int_b^\xi f(x)\,dx + \varphi(b)\int_\xi^b f(x)\,dx$$

$$(a < \xi < b).$$

Обобщенные интегралы. Мы должны теперь уделить некоторое внимание двум важнейшим обобщениям первоначальной идеи интеграла: интегралам с бесконечными пределами и интегралам от неограниченных функций. В обоих случаях речь идет о действительном обобщении самой идеи, а не просто о применении ее в новых условиях; к первоначальной конструкции интеграла присоединяется еще новый, дополнительный пре-

дельный переход; обобщенный интеграл представляет собой уже не предел или грань определенного вида сумм, а предел обыкновенных интегралов.

Вы знаете, что символы $\int\limits_a^{+\infty}$, $\int\limits_{-\infty}^{b}$ и $\int\limits_{-\infty}^{+\infty}$ определяются соответственно как

$$\lim_{b\to+\infty}\int\limits_a^b,\quad \lim_{a\to-\infty}\int\limits_a^b,\quad \lim_{\substack{a\to-\infty\\ b\to+\infty}}\int\limits_a^b.$$

Чтобы лучше рассмотреть основную идею этого обобщения, остановимся на простом частном случае. Пусть функция $f(x)$ — положительная и невозрастающая при $x \geqslant a$ (рис. 28); геометрически интеграл

$$\int\limits_a^b f(x)\,dx$$

иллюстрируется тогда площадью, заштрихованной на рис. 28. При возрастании b эта площадь увеличивается; когда $b \to \infty$, она может либо возрастать безгранично,

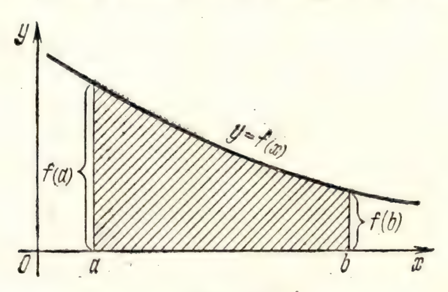


Рис. 28.

либо оставаться ограниченной и, следовательно, стремиться к некоторому пределу. Этот предел мы и называем обобщенным интегралом

$$\int\limits_a^{+\infty} f(x)\,dx;$$

геометрической иллюстрацией его может служить пло-

щадь, заключенная между осью Ox и кривою $y = f(x)$ вправо от точки a; площадь может оставаться конечной, несмотря на то, что соответствующая фигура простирается неограниченно.

Простейший случай второго обобщения — интеграла неограниченной функции — хотя и решает, по-видимому, совсем другую задачу, но по существу почти не отличается от только что рассмотренного. Проще всего в этом убедиться, замечая, что его геометрический образ может быть получен просто зеркальным отражением рис. 28 в биссектрисе угла xOy. Если функция $f(x)$, например, становится неограниченной в соседстве точки a, то

$$\int_a^b f(x)\,dx$$

определяется как предел обыкновенных интегралов $\int\limits_{a+\varepsilon}^{b} f(x)\,dx$ при $\varepsilon \to +0$ (рис. 29). Вы видите, что речь здесь снова идет о возможности приписать определенную величину площади фигуры, простирающейся в бесконечность. Эта фигура — того же типа, как и в первом случае, и только имеет другое расположение; кроме того, вы видите, что в качестве приближенных площадей здесь берутся площади фигур иного типа, чем в первом случае; но все это не имеет, конечно, существенного значения.

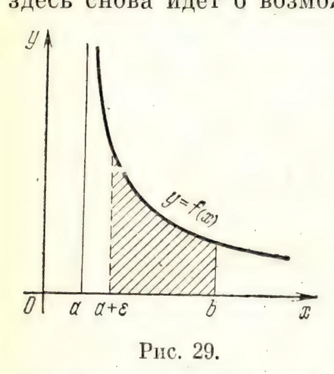


Рис. 29.

В общем случае, т. е. когда поведение функции $f(x)$ какое угодно [1]), определения обобщенных интегралов остаются теми же, хотя и не получают уже столь простой геометрической иллюстрации. В случае, когда со-

[1]) Конечно, в случае неограниченности речь должна идти о неограниченности в соседстве *одной* точки.

ответствующий предел существует, говорят, что данный обобщенный интеграл *существует*, или *имеет смысл*, или *сходится*. Таким образом, вопрос о сходимости обобщенного интеграла того или другого типа всегда представляет собой задачу о существовании предела у некоторой данной функции; все общие предложения теории пределов имеют поэтому силу и для обобщенных интегралов, в частности, имеет место *критерий Коши*, который здесь, очевидно, получает следующую формулировку: *если функция $f(x)$ интегрируема на отрезке $[a, b]$ при любом конечном $b > a$, то для сходимости обобщенного интеграла*

$$\int_a^{+\infty} f(x)\,dx$$

необходимо и достаточно, чтобы, как бы мало ни было $\varepsilon > 0$, при любых достаточно больших b_1 и b_2 мы имели

$$\left|\int_{b_1}^{b_2} f(x)\,dx\right| < \varepsilon$$ [1].

Точно так же имеют место следующие две важные теоремы, являющие собой, как вы узнаете сразу, полную аналогию соответствующим предложениям теории рядов.

1. *Если при $a < x < +\infty$ всегда $0 \leqslant f(x) \leqslant \varphi(x)$, причем $f(x)$ и $\varphi(x)$ интегрируемы на любом отрезке $[a, b]$, где $b > a$, то из сходимости интеграла*

$$\int_a^{+\infty} \varphi(x)\,dx$$

следует сходимость интеграла

$$\int_a^{+\infty} f(x)\,dx.$$

---

[1]) Здесь и в дальнейшем мы ради краткости говорим лишь об одном типе обобщенных интегралов; однако все сказанное с соответствующими изменениями (которые вы, конечно, легко найдете сами) переносится и на все другие типы.

Эта теорема может быть названа «принципом сравнения интегралов» и совершенно аналогична «принципу сравнения рядов» (лекция IV, стр. 99); точно так же и доказательство этой теоремы вы без затруднений проведете самостоятельно на основе изложенной там идеи.

2. *Если функция $f(x)$ интегрируема на любом отрезке $[a, b]$, где $b > a$, то из сходимости интеграла*

$$\int_a^{+\infty} |f(x)|\,dx \tag{11}$$

вытекает сходимость интеграла

$$\int_a^{+\infty} f(x)\,dx. \tag{12}$$

Эта теорема, очевидно, в точности аналогична соответствующей теореме из лекции IV (стр. 105) и может быть доказана в полной аналогии с этой последней; и здесь мы говорим, что интеграл (12) сходится абсолютно, если сходится интеграл (11).

С помощью этих теорем легко устанавливается сходимость большого числа часто встречающихся интегралов; так, из (непосредственно усматриваемой) сходимости интеграла

$$\int_1^{+\infty} \frac{dx}{x^2}$$

в силу теоремы 1 вытекает сходимость интеграла

$$\int_1^{+\infty} \frac{|\sin x|}{x^2}\,dx,$$

а отсюда в силу теоремы 2 — сходимость (и притом абсолютная) интеграла

$$\int_1^{+\infty} \frac{\sin x}{x^2}\,dx.$$

Если в теореме 1 в качестве функции $\varphi(x)$ выбирать различные положительные функции с заведомо сходящимися интегралами, то она дает целый ряд признаков сходимости непосредственно для интегралов с положительной подынтегральной функцией, а с помощью теоремы 2 — и для подынтегральных функций произвольного знака. Несомненно, некоторые из этих признаков вам известны; мы на них останавливаться не будем. Зато мы рассмотрим два признака, менее широко известных и более тонких, так как в них речь идет об интегралах, которые могут и не быть абсолютно сходящимися (откуда, между прочим, видно, что эти признаки не могут быть выведены из теоремы 1).

Признак 1. *Пусть функция* $\varphi(x)$ *монотонна и* $\lim\limits_{x\to\infty}\varphi(x)=0$, *а функция* $f(x)$ *такова, что*

$$\int\limits_a^x f(u)\,du = F(x)$$

остается ограниченным при $a \leqslant x < +\infty$. *Тогда интеграл*

$$\int\limits_a^{+\infty} \varphi(x) f(x)\,dx \tag{13}$$

сходится.

В самом деле, в силу второй теоремы о среднем значении при $a < b_1 < b_2 < +\infty$ существует такая внутренняя точка β отрезка $[b_1, b_2]$, что

$$\int\limits_{b_1}^{b_2} \varphi(x) f(x)\,dx = \varphi(b_1+0)\int\limits_{b_1}^{\beta} f(x)\,dx + \\ + \varphi(b_2-0)\int\limits_{\beta}^{b_2} f(x)\,dx = \varphi(b_1+0)\,[F(\beta)-F(b_1)] + \\ + \varphi(b_2-0)\,[F(b_2)-F(\beta)]. \tag{A}$$

Так как при $b_1 \to +\infty$, $b_2 \to +\infty$ в силу условий теоремы $\varphi(b_1+0)$ и $\varphi(b_2-0)$ стремятся к нулю, в то время как все встречающиеся значения функции $F(x)$

остаются ограниченными, то правая, а следовательно, и левая часть последнего равенства стремятся при этом к нулю, что в силу критерия Коши доказывает сходимость интеграла (13).

Пример. Интеграл

$$\int_1^{+\infty} \frac{\sin x}{x}\, dx \qquad (14)$$

сходится, как непосредственно вытекает из признака 1, если положить $\varphi(x) = x^{-1}$, $f(x) = \sin x$, $F(x) = \cos 1 - \cos x$; можно, однако, легко показать, что сходимость интеграла (14) не является абсолютной.

Признак 2. *Если функция* $\varphi(x)$ *монотонна и ограничена, то из сходимости интеграла* (12) *следует сходимость интеграла* (13).

Для доказательства снова применим вторую теорему о среднем значении; в первом равенстве (A) при $b_1 \to \infty$, $b_2 \to \infty$ величины $\varphi(b_1 + 0)$ и $\varphi(b_2 - 0)$ остаются ограниченными, в то время как оба интеграла стремятся к нулю в силу критерия Коши и предположенной сходимости интеграла (12); значит, левая часть при этом стремится к нулю, откуда опять в силу критерия Коши мы заключаем, что интеграл (13) сходится.

Пример. Полагая $f(x) = \frac{\sin x}{x}$ и $\varphi(x) = \operatorname{arctg} x$, мы из доказанной сходимости интеграла (14) заключаем, что сходится и интеграл

$$\int_1^{+\infty} \frac{\sin x \operatorname{arctg} x\, dx}{x}.$$

Двойные интегралы. Тот специальный процесс суммирования, который лежит в основании интегрального исчисления, может быть с успехом применен и к функциям, зависящим от нескольких переменных. В большом числе прикладных областей, и в особенности в механике и физике, такие многомерные или, как их чаще называют, «кратные» интегралы играют значительную роль. Теория этих интегралов, не содержащая в прин-

ципиальном отношении почти ничего нового сравнительно с теорией обыкновенного интеграла, с формальной стороны очень громоздка.

В дальнейшем мы, оставаясь в пределах двумерного случая, кратко покажем, как идеи, изложенные нами в начале настоящей лекции, в почти неизменном виде находят себе применение к определению кратных интегралов и установлению целого ряда их свойств.

Пусть мы имеем функцию двух переменных $z = f(x, y)$, определенную и ограниченную в некоторой замкнутой [1]) ограниченной области D координатной плоскости xy. Пусть M и m означают соответственно верхнюю и нижнюю грани функции $f(x, y)$ в области D, площадь которой мы условимся обозначать той же буквой D. Подобно тому как мы поступали в одномерном случае, мы и здесь будем рассматривать различные разбиения T области D на части. Конечно, здесь мы имеем картину значительно более сложную, чем раньше. В одномерном случае как основная область, так и те части, на которые мы ее разбивали, имели одну форму — это были отрезки; в том случае, который мы рассматриваем теперь, как область D, так и те части $\Delta_1, \Delta_2, \ldots, \Delta_n$, на которые мы ее разбиваем, могут иметь весьма различную форму. Для теории выгодно по возможности ничем не ограничивать вида всех этих областей. Важно, конечно, только, чтобы каждая из этих областей имела определенную площадь и чтобы площадь общей части (т. е. взаимной границы) любых двух различных областей Δ_i и Δ_k была равна нулю. В остальном мы не будем пока налагать на разбиение T никаких специальных требований.

Пусть, как в одномерном случае, M_k и m_k соответственно означают верхнюю и нижнюю грани функции $f(x, y)$ на части Δ_k; положим, как прежде,

$$S_T = \sum_{k=1}^{n} M_k \Delta_k, \quad s_T = \sum_{k=1}^{n} m_k \Delta_k,$$

где Δ_k символизирует, конечно, площадь области, обозначенной тем же символом. Как и ранее, мы имеем,

[1]) То есть точки границы присоединяются к области.

очевидно,

$$mD \leqslant s_T \leqslant S_T \leqslant MD,$$

так что множества всех сумм S_T и s_T ограничены сверху и снизу. Пусть $\overline{I}$ и $\underline{I}$ соответственно означают нижнюю грань множества всех сумм S_T и верхнюю грань множества всех сумм s_T; мы и здесь будем называть эти два числа соответственно верхним и нижним интегралами функции $f(x, y)$ в области D.

Если $\overline{I} = \underline{I}$, то функция $f(x, y)$ *интегрируема* в области D и интеграл ее равен

$$\overline{I} = \underline{I} = I = \iint\limits_D f(x, y)\, dx\, dy.$$

В основе всей теории таких «двойных» интегралов лежит ряд предложений, в точности аналогичных тем, которые мы имели в одномерном случае. Мы теперь кратко рассмотрим их, останавливаясь на доказательствах только в тех случаях, где они отличаются от данных нами прежде.

Условимся за отсутствием более удобного термина называть *ячейками* те части $\Delta_1, \Delta_2, \ldots, \Delta_n$, на которые мы разбиваем основную область D. *Диаметром* данной ячейки мы будем называть верхнюю грань расстояний между всевозможными парами принадлежащих ей точек. Для каждого разбиения T обозначим через d_T наибольший из диаметров ячеек $\Delta_1, \Delta_2, \ldots, \Delta_n$; очевидно, эта величина должна играть здесь ту же роль, как величина l_T в одномерном случае. Будем, наконец, говорить, что разбиение T' есть *продолжение* разбиения T, если каждая ячейка разбиения T' принадлежит целиком одной из ячеек разбиения T.

Теперь установим четыре вспомогательных предложения, в точности совпадающих с теми четырьмя леммами, которые мы имели в одномерном случае.

Лемма I. *Если разбиение T' есть продолжение разбиения T, то*

$$S_{T'} \leqslant S_T, \quad s_{T'} \geqslant s_T.$$

Доказательство может быть буквально перенесено из одномерного случая.

Лемма II. *Для любых двух разбиений T_1 и T_2*

$$S_{T_1} \geqslant s_{T_2}.$$

Доказательство, указанное нами для одномерного случая, может быть в точности повторено. Пояснения требует лишь построение разбиения T, служащего продолжением каждого из двух данных разбиений T_i; мы просто принимаем за ячейку разбиения T совокупность всех точек, принадлежащих одной из ячеек разбиения T_1 и одной из ячеек разбиения T_2. Составляя всевозможные такие пары пересекающихся ячеек, мы и получим все ячейки разбиения T; по самому определению продолжения тогда очевидно, что разбиение T является одновременно продолжением как разбиения T_1, так и разбиения T_2.

Лемма III. $\overline{I} \geqslant \underline{I}$.

Это предложение, как и в одномерном случае, есть непосредственное следствие леммы II.

Лемма IV. *Как бы мало ни было $\varepsilon > 0$, существует такое число $\delta > 0$, что при любом разбиении T, удовлетворяющем требованию $d_T < \delta$,*

$$S_T < \overline{I} + \varepsilon, \quad s_T > \underline{I} - \varepsilon.$$

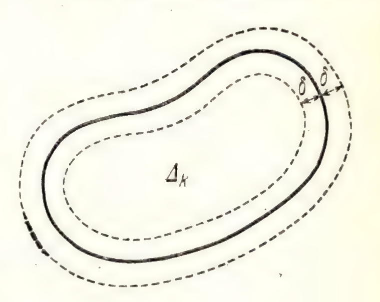


Рис. 30.

Короче говоря, при $d \to 0$ суммы S_T стремятся к своей нижней, а суммы s_T — к своей верхней грани.

Доказательство этого предложения принципиально также не отличается от того, которое мы привели для одномерного случая; однако здесь более сложное строение ячеек все же требует, для соблюдения формальной безупречности, детального изложения.

Прежде всего, так как $\overline{I}$ определяется как нижняя грань всех сумм S_T, то существует разбиение T_0, для которого $S_{T_0} < \overline{I} + \frac{\varepsilon}{2}$. Пусть Δ_k — любая ячейка этого

разбиения T_0. Совокупность точек плоскости, отстоящих от контура этой ячейки не более чем на δ, образует, очевидно, род «кольца» вокруг этого контура (на рис. 30 границы этого «кольца» намечены пунктиром). Нетрудно убедиться, что площадь такого «кольца» не превосходит $2\delta l_k$, где l_k — длина контура ячейки Δ_k [1]). Положим

$$\sum_{k=1}^{n} l_k = L;$$

тогда площадь области D_1, образованной совокупностью всех «колец» построенного типа, будет не больше чем $2\delta L$ (она может быть меньше, так как «кольца» могут перекрываться между собой). Пусть теперь T — любое разбиение, для которого $d_T < \delta$; ячейки этого разбиения мы разделим на две группы, относя к первой каждую ячейку, целиком принадлежащую области D_1, а ко второй — все другие ячейки; соответственные этим двум группам части суммы S_T обозначим через $S_T^{(1)}$ и $S_T^{(2)}$. Оценим каждую из этих сумм в отдельности. При этом, как и в одномерном случае, мы, не ограничивая общности рассуждения, допустим, что $f(x, y) \geqslant 0$ во всей области D.

Так как в членах суммы $S_T^{(1)}$ первые множители не превосходят M, а сумма площадей ячеек, участвующих в $S_T^{(1)}$, не превосходит площади области D_1 (которая, как мы видели, не больше чем $2\delta L$), то

$$S_T^{(1)} \leqslant 2ML\delta; \tag{15}$$

что же касается ячеек, входящих в состав суммы $S_T^{(2)}$, то каждая из них целиком принадлежит какой-либо одной из ячеек Δ_k разбиения T_0; в самом деле, ячейка Δ второй группы, не удовлетворяющая этому требованию, должна была бы содержать в себе точку P контура

---

[1]) Само собой разумеется, что здесь мы предъявляем уже негласно к ячейкам наших разбиений более стеснительные требования, чем вначале; так, мы постулируем, что каждая ячейка есть односвязная область, имеющая контур конечной длины. К сожалению, в рамках наших лекций мы лишены возможности подробнее осветить эти детали.

какой-либо из ячеек Δ_k; но, будучи ячейкой второй группы, Δ не может целиком принадлежать «кольцу», окружающему этот контур, а потому должна содержать точку Q, лежащую вне этого «кольца»; однако тогда расстояние точек P и Q, а значит и подавно диаметр ячейки Δ превосходит δ, что невозможно.

Мы не уменьшим сумму $S_T^{(2)}$, если в каждом ее члене заменим первый множитель величиной M_k, соответствующей той ячейке Δ_k разбиения T_0, часть которой составляет входящая в данный член ячейка Δ (второй группы) разбиения T; с другой стороны, не уменьшенная, таким образом, сумма в силу неотрицательности функции $f(x, y)$ будет, очевидно, все же не больше чем $S_{T_0}^{(2)}$; таким образом,

$$S_T^{(2)} \leqslant S_{T_0} < \bar{I} + \frac{\varepsilon}{2}.$$

Сопоставляя это неравенство с неравенством (15) и выбирая число δ, до сих пор оставшееся произвольным, меньшим, чем $\frac{\varepsilon}{4ML}$, мы получаем

$$S_T = S_T^{(1)} + S_T^{(2)} < 2ML\delta + \bar{I} + \frac{\varepsilon}{2} < \bar{I} + \varepsilon.$$

В точности таким же способом можно показать, что $s_T > I - \varepsilon$ при условии $d_T < \delta$, чем и завершается доказательство леммы IV.

Теорема. *Если функция $f(x, y)$ интегрируема в области D, то для всех разбиений T с достаточно малым d_T*

$$\left| \sum_{k=1}^{n} f(\xi_k, \eta_k)\, \Delta_k - I \right| < \varepsilon,$$

где ε — любое заранее выбранное положительное число, а (ξ_k, η_k) $(k = 1, 2, \ldots, n)$ — любая точка ячейки Δ_k.

Говоря короче, сумма $\sum\limits_{k=1}^{n} f(\xi_k, \eta_k)\, \Delta_k$ стремится к интегралу I при $d_T \to 0$ равномерно относительно всех возможных выборов как разбиения T, так и точек (ξ_k, η_k).

Обозначая для краткости эту сумму просто через Σ_T, мы из очевидных неравенств

$$m_k \leqslant f(\xi_k, \eta_k) \leqslant M_k$$

имеем при любом разбиении и любом выборе точек (ξ_k, η_k)

$$s_T \leqslant \sum_T \leqslant S_T,$$

а значит, в силу леммы IV при достаточно малом d_T

$$I - \frac{\varepsilon}{2} \leqslant \sum_T \leqslant I + \frac{\varepsilon}{2},$$

откуда $|\Sigma_T - I| < \varepsilon$. Обращение доказанной теоремы также справедливо и доказывается в точности так же, как в одномерном случае.

Универсальный необходимый и достаточный признак сходимости

$$S_T - s_T = \sum_{k=1}^{n} \omega_k \Delta_k \to 0 \qquad (d_T \to 0)$$

доказывается, как прежде: 1) если этот признак выполнен, то в силу неравенств $S_T \geqslant \overline{I} \geqslant \underline{I} \geqslant s_T$ (при любом разбиении) мы при достаточно малом d_T имеем $\overline{I} - \underline{I} \leqslant S_T - s_T < \varepsilon$, откуда $\overline{I} = \underline{I}$ ввиду произвольности ε; 2) если функция $f(x, y)$ интегрируема, то в силу леммы IV при достаточно малом d_T мы имеем $S_T < I + \varepsilon$, $s_T > I - \varepsilon$, откуда $S_T - s_T < 2\varepsilon$, и, значит, $S_T - s_T \to 0$ при $d_T \to 0$.

Наконец, интегрируемость непрерывных функций доказывается здесь в точности тем же рассуждением, как и в одномерном случае; в основании этого доказательства и здесь лежит тот факт, что непрерывная функция всегда обладает равномерной непрерывностью (в закрытой, т. е. содержащей свою границу, области).

Вычисление двойных интегралов. Уже по поводу обыкновенных интегралов мы выше имели случай говорить о том, что процесс суммирования, лежащий в основании определения интеграла, почти ничего не дает, если мы пытаемся сделать его орудием фактического вычисления интегралов; тем менее мы можем, конечно, рассчитывать, чтобы двойные интегралы удоб-

но было вычислять с помощью только что описанного процесса двумерного суммирования.

В случае обыкновенных интегралов мы видели, что мощные общие методы вычисления интегралов рождаются только путем использования связи интегрального исчисления с дифференциальным. Для двойных (и вообще многомерных) интегралов такие связи тоже могут быть обнаружены; однако мы получаем наиболее общий и эффективный способ вычисления двойного интеграла известным приведением этой задачи к двум последовательным одномерным интеграциям; рассмотрим кратко, как это делается.

Пусть область D, в которой определена непрерывная функция $f(x, y)$, такова, что всякая прямая, параллельная одной из координатных осей, пересекает ее контур

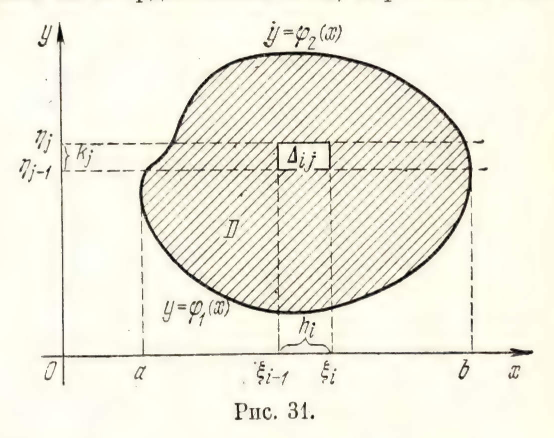


Рис. 31.

не более чем в двух точках (рис. 31)[1]). Разбиение T этой области, которое мы производим для построения интеграла I функции $f(x, y)$, может иметь ячейки какой угодно формы, лишь бы диаметры их становились сколь угодно малыми. Мы теперь проведем это разбиение с помощью сети прямых, параллельных координатным осям, так что ячейки будут прямоугольниками (за

---

[1]) Автор накладывает на область D излишние ограничения.— *Прим. ред.*

исключением тех, которые расположены в непосредственной близости контура области D). Мы будем обозначать, считая слева направо, через $\xi_1, \xi_2, \ldots$ абсциссы прямых, параллельных оси Oy; через $\eta_1, \eta_2 \ldots$ (считая снизу вверх) — ординаты прямых, параллельных оси Ox; через a и b — соответственно нижнюю и верхнюю грани абсцисс точек области D; через $y = \varphi_1(x)$ и $y = \varphi_2(x)$ — соответственно уравнения нижней и верхней границ области D; через Δ_{ij} — ячейку, ограниченную прямыми $x = \xi_{i-1}$, $x = \xi_j$, $y = \eta_{j-1}$, $y = \eta_j$; положим, наконец, $\xi_i - \xi_{i-1} = h_i$, $\eta_i - \eta_{i-1} = k_j$, так что $\Delta_{ij} = h_i k_j$, если только ячейка Δ_{ij} имеет прямоугольную форму, т. е. не лежит у контура области D.

Если все h_i, k_j достаточно малы, то d_T сколь угодно мало и интеграл I сколь угодно мало отличается от суммы

$$S = \sum_{i,j} f(\xi_j, \eta_j)\, \Delta_{ij}. \tag{16}$$

Освободимся прежде всего от неудобных ячеек непрямоугольной формы, расположенных у контура области D. Обозначим через δ наибольшее из чисел h_i, k_j. Полоса, заключенная между прямыми $x = \xi_{i-1}$ и $x = \xi_i$, может содержать ячейки непрямоугольной формы вверху и внизу; на рис. 32 изображена схематически примерная форма верхнего окончания такой полосы; ячейки (на рис. 32 их три) заштрихованы и заключены в прямоугольник, основание которого равно h_i, а высота, очевидно, не превосходит $\omega_i + 2\delta$, где ω_i — колебание функции $\varphi_2(x)$ на отрезке $[\xi_{i-1}, \xi_i]$ (этот прямоугольник и изображен на рис. 32); если δ достаточно мало, то в силу равномерной непрерывности функции $\varphi_2(x)$ все ω_i, а значит, и все $\omega_i + 2\delta$ будут меньше как угодно малого положительного числа ε, так что сумма площадей изображенных на рис. 32 непря-

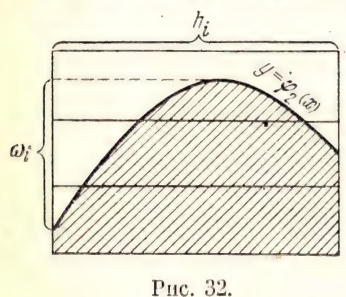


Рис. 32.

моугольных ячеек будет меньше, чем εh_i, а следовательно, сумма площадей всех таких ячеек, прилегающих к верхней половине контура области D, меньше чем $\varepsilon \sum_i h_i = \varepsilon (b - a)$; такую же оценку получает, очевидно, и площадь множества ячеек, примыкающих к нижней половине этого контура, так что все вместе взятые «неполные» ячейки дают площадь, не превышающую $2\varepsilon (b - a)$; следовательно, распространенная на эти ячейки часть суммы (16) по абсолютной величине не превосходит $2M(b-a)\varepsilon$ (где M — верхняя грань функции $|f(x, y)|$ в области D) и потому сколь угодно мала при достаточно малом δ. Поэтому, распространяя во всем дальнейшем сумму (16) только на прямоугольные ячейки, мы по-прежнему будем иметь $|I - \widehat{S}| < \varepsilon$ для всех разбиений с достаточно малым d_T; при этом теперь

$$\widehat{S} = \sum_{i,j} f(\xi_i, \eta_j) h_i k_j = \sum_i h_i \sum_j f(\xi_i, \eta_j) k_j.$$

Если мы, закрепив определенную систему значений ξ_1, ξ_2, ..., будем безгранично уменьшать все разности $k_j = \eta_j - \eta_{j-1}$, то внутренняя сумма $\sum_j f(\xi_i, \eta_j) k_j$ будет иметь своим пределом интеграл

$$\int_{\varphi_1(\xi_i)}^{\varphi_2(\xi_i)} f(\xi_i, y)\, dy,$$

который мы ради краткости обозначим через $F(\xi_i)$, полагая вообще

$$F(x) = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\, dy$$

(при интеграции x считается постоянным, так что $f(x, y)$ интегрируются как функция одной переменной y). Мы можем взять величины k_j столь малыми, что для любого i

$$\left| \sum_j f(\xi_i, \eta_j) k_j - F(\xi_i) \right| < \varepsilon,$$

откуда

$$\left| h_i \sum_j f(\xi_i, \eta_j) k_j - F(\xi_i) h_i \right| < \varepsilon h_i$$

и

$$\left| \sum_i h_i \sum_j f(\xi_i, \eta_j) k_j - \sum_i F(\xi_i) h_i \right| =$$
$$= \left| \widehat{S} - \sum_i F(\xi_i) h_i \right| < \varepsilon \sum_i h_i = \varepsilon (b-a);$$

наконец, при достаточно малых h_i сумма $\sum\limits_i F(\xi_i) h_i$, очевидно, сколь угодно близка к интегралу

$$\int_a^b F(x)\,dx.$$

Итак, мы видим, что при надлежащем выбранном разбиении области D все три разности: $I - \widehat{S}, \widehat{S} - \sum\limits_i F(\xi_i) h_i$ и $\sum\limits_i F(\xi_i) h_i - \int\limits_a^b F(x)\,dx$ — становятся угодно малыми; значит, то же справедливо и для их суммы $I - \int\limits_a^b F(x)\,dx$, которая, не завися ни от каких разбиений, должна поэтому равняться нулю. Таким образом, мы получаем

$$\iint\limits_D f(x, y)\,dx\,dy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)\,dy,$$

так что вычисление двойного интеграла действительно сводится к двум последовательным обыкновенным интеграциям. Внутренний интеграл, разумеется, зависит от x, но уже не зависит от y. Конечно, вместо выбранного нами порядка интеграции может быть взят и обратный; это замечание иногда получает существенное практическое значение: именно, может случиться, что, меняя порядок интеграции, мы придем к функциям, значительно легче интегрируемым.

Общая идея интеграла. Наша лекция затянулась. И все же, прежде чем закончить ее, мы хотим еще

бросить общий взгляд на картину, лежащую в основе всякого процесса интеграции, будет ли то обыкновенный, или многомерный, или даже еще более общий и отвлеченный интеграл.

Прежде всего, мы имеем дело с определенной областью D некоторого пространства, причем этот термин «пространство» надо понимать здесь в весьма широком значении: это может быть и прямая, и плоскость, и наше обычное трехмерное, и любое многомерное, и даже какое-либо совсем иное пространство; мы не собираемся здесь давать этому термину общего определения; для нашей цели важно только, чтобы, во-первых, две любые точки нашего пространства были отделены определенным расстоянием и чтобы, во-вторых, область D и те ее части («ячейки»), которые мы будем рассматривать, имели определенную *протяженность*; эта протяженнность, смотря по характеру пространства, называется обычно длиной, площадью, обьемом и т. п.; для общности мы будем называть ее просто *мерой* соответствующей части нашего пространства.

Теперь представьте себе, что в области D распределена некоторая субстанция; пусть это слово не страшит вас — мы не сделаем попытки *определять* выражаемое им понятие; для нас субстанция есть просто *нечто такое, чего приходится определенная порция на всякую часть области D*; это может быть масса, электрический заряд, теплота — все величины, могущие быть самым различным образом распределенными в данной области D; это может быть, например, количество осадков, выпадающих на ту или иную часть плоской области D за данный промежуток времени. Само собой разумеется, что для построения нашей математической картины природа этой субстанции глубоко безразлична, нам важно только, чтобы на каждую рассматриваемую нами часть Δ области D приходилось определенное количество $F(\Delta)$ этой субстанции. При каждом разбиении T области D на части («ячейки») $\Delta_1, \Delta_2, \ldots, \Delta_n$ общее количество субстанции $F(D)$ слагается из количеств субстанции, приходящихся на отдельные ячейки:

$$F(D) = F(\Delta_1) + F(\Delta_2) + \ldots + F(\Delta_n).$$

Во всех реальных интерпретациях нарисованной нами формальной схемы играет существенную роль понятие «плотности» субстанции в данной точке P области D, с математической стороны определяемое, как мы сейчас увидим, *дифференцирующим* процессом. Если Δ — любая часть области D, имеющая определенную меру $m\Delta$, то естественно называть отношение $\frac{F(\Delta)}{m\Delta}$ *средней плотностью* нашей субстанции в области Δ; это — количество субстанции, в среднем приходящееся в области Δ на единицу протяженности (меры). Пусть теперь P — любая точка области D; если мы окружим эту точку какой-либо областью Δ малого диаметра $d(\Delta)$ (а диаметр любой области в нашем пространстве вполне определен, ибо определено расстояние между двумя любыми точками), то средняя плотность субстанции в этой малой области Δ будет характеризовать нам густоту распределения этой субстанции в ближайшем соседстве точки P; а если эта средняя плотность, как мы допустим, стремится к определенному пределу $f(P)$, когда область Δ стягивается к точке P ($d(\Delta) \to 0$), то этот предел $f(P)$ мы называем плотностью нашей субстанции в точке P. Именно так определяется плотность для всех физических субстанций — массы, электричества и т. п. Мы видим, что плотность субстанции $f(P)$ есть функция *точки* нашего пространства, в то время как количество субстанции $F(\Delta)$ является функцией *области*. Можно сказать, что функция $F(\Delta)$ дает *интегральную* (по областям), а функция $f(P)$ — *локальную* (по точкам) характеристику распределения нашей субстанции в области D. Если функция $F(\Delta)$ задана, то $f(P)$ находится из нее с помощью некоторого *дифференцирующего* процесса; сущность дифференцирующего процесса в принципиальном плане, пожалуй, лучше прямо и определять как переход от интегрального описания какого-либо явления к локальной его характеризации.

Но пусть теперь, обратно, нам задана в каждой точке P области D плотность распределения $f(P)$ некоторой субстанции и требуется определить общее количество $F(D)$ субстанции, заключенной в этой области.

Эта задача и есть проблема интегрирования в ее наиболее общей форме: переход от локальной к интегральной характеризации изучаемого явления. Пусть для простоты функция $f(P)$ непрерывна (а значит, и равномерно непрерывна) в области D (мы предполагаем эту область замкнутой и ограниченной). Для решения нашей задачи мы тогда разбиваем область D на ячейки $\Delta_1, \Delta_2, \ldots, \Delta_n$ малых диаметров; в каждой ячейке Δ_k выбираем произвольную точку P_k; $f(P_k)$ есть предел отношения $\frac{F(\Delta)}{m\Delta}$ при условии, что область Δ, все время заключая в себе точку P_k, безгранично уменьшается в диаметре; поэтому, если ячейки Δ_k очень малы, то

$$\left| f(P_k) - \frac{F(\Delta_k)}{m\Delta_k} \right| < \varepsilon,$$

где ε — любое заданное заранее положительное число; отсюда

$$F(\Delta_k) - \varepsilon m\Delta_k < f(P_k)\, m\Delta_k < F(\Delta_k) + \varepsilon m\Delta_k,$$

а суммируя эти неравенства по k, находим

$$F(D) - \varepsilon mD < \sum_{k=1}^{n} f(P_k)\, m\Delta_k < F(D) + \varepsilon mD,$$

откуда в пределе при неограниченно убывающих диаметрах ячеек

$$\lim \sum_{k=1}^{n} f(P_k)\, m\Delta_k = F(D).$$

Таким образом, мы видим, что по данной функции $f(P)$ (плотности распределения) величина $F(D)$ (т. е. общее количество субстанции) действительно определяется с помощью хорошо знакомого нам интегрирующего процесса: мелкое разбиение области D на ячейки Δ_k, выбор точек P_k, составление суммы произведений вида $f(P_k)\, m\Delta_k$ по всем ячейкам и, наконец, переход к пределу при условии, что диаметры ячеек стремятся к нулю.

Таким образом, с рассматриваемой нами весьма общей точки зрения процесс интеграции в каких угодно

пространствах представляется методом, позволяющим определить приходящееся на ту или иную область количество субстанции, зная плотность распределения этой субстанции в каждой отдельной точке рассматриваемой основной области. Всякий такой метод имеет своим неизбежным коррелятом некоторый (решающий обратную задачу) дифференцирующий процесс, целью которого является нахождение локальной плотности распределенной субстанции, когда известно количество ее, приходящееся на любую область. Имеется полная возможность, достаточно точно перечислив основные предпосылки, строить теорию интеграции на этой общей абстрактной базе; тогда обыкновенные, двойные и другие специальные интегралы становятся частными случаями такой общей теории, важнейшие свойства которых раз навсегда установлены этой общей теорией и не требуют особых для каждого отдельного случая доказательств.

Л Е К Ц И Я VII

РАЗЛОЖЕНИЕ ФУНКЦИЙ В РЯДЫ

Ряды в роли аппарата для исследования функций.— Разложение в степенные ряды.— Ряды многочленов. Теорема Вейерштрасса.— Тригонометрические ряды.— Коэффициенты Фурье.— Аппроксимация в среднем.— Замкнутость системы тригонометрических функций.— Сходимость рядов Фурье для функций, имеющих ограниченную интегрируемую производную.— Распространение на произвольные отрезки.

Ряды в роли аппарата для исследования функций. Когда мы даем синусу и косинусу их первоначальное геометрическое определение, то целый ряд важнейших свойств этих функций мы выводим непосредственно из их определения. Когда мы затем вводим для этих функций общепринятые обозначения $\sin x$ и $\cos x$, то эти «аналитические выражения» ничего нового к нашим сведениям о выражаемых ими функциях, разумеется, не прибавляют, точно так же как мы не узнаем ничего нового о свойствах функции Дирихле, обозначая ее через $f(x)$. Но когда мы позднее находим для тригонометрических функций новые аналитические выражения, представляющие их в виде степенных рядов:

$$\sin x = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!}, \quad \cos x = \sum_{n=0}^{\infty} (-1)^{n} \frac{x^{2n}}{(2n)!},$$

то с помощью этих представлений мы легко обнаруживаем целый ряд таких свойств этих функций, которые мы совсем не могли бы или могли бы лишь очень сложным путем получить из их первоначальных определений; и прежде всего, эти разложения дают нам

возможность с максимальной принципиальной простотой вычислять значения тригонометрических функций при любых значениях аргумента.

В лекции III мы говорили о тех опасностях, которые таит в себе чрезмерное преклонение перед аналитическими выражениями; действительно, в среде представителей прикладных наук мы часто встречаемся с вредными последствиями этого явления, о которых мы в свое время подробно говорили; но нельзя отрицать, что среди современных математиков мы иногда встречаем противоположную крайность, способную принести не меньше вреда: принципиальное пренебрежение к аналитическим выражениям, практически ведущее к неумению и беспомощности в обращении с ними.

Аналитическое выражение, если мы с самого начала подойдем к нему в роли хозяев, а не покорных и немудрствующих подчиненных, если мы будем видеть в нем лишь то, что оно есть на самом деле,— послушное нашим целям и замыслам орудие исследования функций,— может сыграть в этом исследовании огромную, решающую роль и принести чрезвычайно много пользы. Более того, можно считать нормальным и естественным такое положение вещей, когда математик, стоящий перед задачей исследования какой-либо определенной функции, прежде всего ищет для нее удобное аналитическое выражение; из деталей такого выражающего данную функцию аналитического аппарата в большинстве случаев и удается наиболее рациональным способом вычитать нужные свойства этой функции.

Среди различных аналитических аппаратов, могущих в этом смысле служить орудиями исследования функций, первое место по своей простоте, гибкости, прозрачности и удобству употребления, без сомнения, занимают функциональные ряды. Идея этого важнейшего аналитического аппарата очень проста: функция, которую мы хотим исследовать, представляется как предел последовательности других, более простых и доступных исследованию функций (частичных сумм изображающего ряда); если такая частичная сумма на всем рассматриваемом участке весьма близка к изучае-

мой функции, то мы имеем основание рассчитывать из свойств этой приближенной частичной суммы узнать, хотя бы приближенно, некоторые свойства изучаемой функции; в частности, умея приближенно вычислять, при тех или других значениях аргумента, значения этих частичных сумм, мы тем самым имеем метод для приближенного вычисления соответствующих значений изучаемой функции.

Какими же функциями всего удобнее и выгоднее пользоваться в качестве элементов разложения, т. е. членов того ряда, который призван изображать собою изучаемую функцию и помочь нам в ее исследовании? На этот вопрос, как и естественно ожидать, нельзя дать единого ответа, пригодного для всех случаев; здесь почти все зависит от природы изучаемой функции и характера тех задач, которые мы по отношению к ней себе ставим. Необходимо только отметить, что существует несколько важнейших типов функциональных рядов, в такой степени зарекомендовавших себя в этой своей роли, что применение их совершается на каждом шагу и, естественно, вызвало создание соответствующих общих теорий; сюда относятся в первую очередь *степенные ряды* (где элементами разложения служат целые, в первую очередь — неотрицательные степени независимой переменной) и *тригонометрические ряды* (с элементами вида $\sin kx$, $\cos kx$, где $k = 0, 1, 2, \ldots$). Однако в очень многих случаях в качестве элементов разложения бывает удобнее выбрать не эти простейшие, «универсальные», а совсем другие функции, хотя и не столь простые, но зато по своим свойствам более тесно связанные с изучаемой функцией (например, так называемые «собственные функции» в краевых задачах). Вообще при выборе типа разложения основным руководящим принципом должно служить отсутствие всякой предвзятости, которое в максимальной степени позволило бы учесть в каждом отдельном случае всю специфику поставленных нами перед собою задач. В дальнейшем мы кратко коснемся лишь наиболее важных в принципиальном отношении вопросов, связанных с разложением функций в степенные и тригонометрические ряды.

Разложение в степенные ряды. Мы знаем, что областью сходимости степенного ряда всегда служит некоторый (открытый, закрытый или полуоткрытый) отрезок с концами $-r$ и r. Какими свойствами должна обладать функция $f(x)$ для того, чтобы она допускала разложение

$$f(x) = \sum_{n=0}^{\infty} a_n x^n,$$

сходящееся на этом отрезке? Мы знаем, что для этого функция $f(x)$ должна быть непрерывной на открытом отрезке $(-r, r)$ (лекция IV, стр. 124); однако этого далеко не достаточно; покажем прежде всего, что при $-r < x < r$ функция $f(x)$ должна иметь производную в точке x; одновременно мы установим, что эта производная $f'(x)$ изображается степенным рядом

$$\sum_{n=1}^{\infty} n a_n x^{n-1}, \tag{1}$$

получаемым почленным дифференцированием данного ряда и сходящимся на открытом отрезке $(-r, r)$: при доказательстве надо все время иметь в виду, что ни существование производной $f'(x)$, ни сходимость ряда (1) нам заранее не даны, так что оба факта должны быть установлены в процессе рассуждения.

Пусть $|x| < r$, и пусть ρ — любое число, заключенное между $|x|$ и r; при $|h| < \rho - |x|$ мы имеем $|x+h| < \rho < r$ и, следовательно,

$$f(x+h) = \sum_{n=0}^{\infty} a_n (x+h)^n,$$

откуда

$$\frac{f(x+h) - f(x)}{h} = \sum_{n=0}^{\infty} a_n \frac{(x+h)^n - x^n}{h} = S_N(h) + R_N(h),$$

где положено

$$S_N(h) = \sum_{n=0}^{N} a_n \frac{(x+h)^n - x^n}{h},$$

$$R_N(h) = \sum_{n=N+1}^{\infty} a_n \frac{(x+h)^n - x^n}{h}.$$

Так как

$$\left|\frac{(x+h)^n-x^n}{(x+h)-x}\right| = \left|(x+h)^{n-1}+(x+h)^{n-2}x+\ldots\right.$$
$$\left.\ldots+x^{n-1}\right| < \rho^{n-1}+\rho^{n-2}|x|+\ldots+|x|^{n-1} =$$
$$= \frac{\rho^n-|x|^n}{\rho-|x|} < \frac{\rho^n}{\rho-|x|},$$

то при $|h| < \rho-|x|$

$$|R_N(h)| < \frac{1}{\rho-|x|}\sum_{n=N+1}^{\infty}|a_n|\rho^n;$$

но $\rho < r$, и, следовательно, сумма в правой части этого неравенства, будучи остатком сходящегося ряда, при достаточно большом N становится меньше как угодно малого положительного числа ε; таким образом, если N достаточно велико и $|h| < \rho-|x|$, то

$$|R_N(h)| < \varepsilon,$$

и, следовательно,

$$S_N(h)-\varepsilon < \frac{f(x+h)-f(x)}{h} < S_N(h)+\varepsilon.$$

Если мы теперь, оставляя N неизменным, заставим h стремится к нулю, то сумма $S_N(h)$, очевидно, будет иметь пределом сумму

$$S_N = \sum_{n=1}^{N} na_nx^{n-1};$$

следовательно, последние неравенства позволяют утверждать, что как верхний, так и нижний пределы величины $\frac{f(x+h)-f(x)}{h}$ при этом будут не более чем на ε отличаться от величины S_N и, следовательно, не более чем на 2ε отличаться друг от друга; в силу произвольности ε отсюда вытекает, что эти пределы совпадают между собой, т. е. что

$$\lim_{h\to\infty}\frac{f(x+h)-f(x)}{h} = f'(x)$$

существует и не больше чем на ε отличается от

суммы S_N:

$$|f'(x) - S_N| \leqslant \varepsilon,$$

при единственном условии, что N достаточно велико; но это означает, что ряд (1) сходится и имеет своей суммой $f'(x)$. Таким образом, мы доказали все наши утверждения.

Итак, функция, изображаемая степенным рядом, должна быть не только непрерывной, но и дифференцируемой. Но и этого мало. Мы только что видели, что $f'(x)$ изображается степенным рядом, сходящимся для всех $|x| < r$; в силу только что доказанной теоремы поэтому всюду на открытом отрезке $(-r, r)$ должна существовать и вторая производная $f''(x)$; продолжая же это рассуждение, мы приходим к выводу, что *функция, изображающаяся на некотором отрезке степенным рядом, должна в каждой внутренней точке этого отрезка иметь производные всех порядков.* При этом каждая из этих производных изображается на том же (открытом) отрезке степенным рядом, получаемым из данного ряда путем повторенного соответственное число раз почленного дифференцирования, так что

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1},$$

и вообще

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\ldots(n-k+1)\, a_n x^{n-k}.$$

Полагая в этом соотношении $x = 0$, мы находим

$$f^{(k)}(0) = k!\, a_k,$$

откуда

$$a_k = \frac{f^{(k)}(0)}{k!} \quad (k = 0, 1, 2, \ldots).$$

Этим мы одновременно *доказали единственность* разложения функции в степенной ряд и нашли выражение коэффициентов этого разложения через значения производных данной функции при $x = 0$. Таким образом, если функция $f(x)$ вообще разлагается в

степенной ряд, то это разложение имеет вид

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n. \qquad (2)$$

Это — так называемый *ряд Маклорена*; полагая $x = a + h$, $f(a+h) = \varphi(h)$, мы получаем

$$\varphi(h) = \sum_{n=0}^{\infty} \frac{\varphi^{(n)}(0)}{n!} h^n,$$

что по возвращении к старым обозначениям приводит к более общему *ряду Тэйлора:*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n. \qquad (3)$$

Положение вещей становится особенно ясным при сопоставлении всех этих фактов с тем, что мы говорили в лекции V о формулах Тэйлора и Маклорена. Там у нас не было речи о бесконечных рядах: величину

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

мы называли остаточным членом данной формулы Маклорена и исследовали ее поведение, предполагая x бесконечно малым. Теперь мы видим, что поведением этой же величины решается вопрос о разложимости функции $f(x)$ в степенной ряд; но в лекции V мы считали n постоянным и заставляли x стремиться к нулю; теперь же, напротив, мы должны, выбрав какое-либо значение x, закрепить его и заставить число n безгранично возрастать. Условие

$$\lim_{n\to\infty} R_n(x) = 0,$$

очевидно, необходимо и достаточно для того, чтобы имела место формула (2). Обычно разложимость данной функции в ряд Маклорена устанавливается вы-

бором для остаточного члена $R_n(x)$ какой-либо из большого числа найденных для него форм, с несколькими из которых мы ознакомились в лекции V. Какую из этих форм выгоднее всего выбрать для этой цели,— это полностью зависит, конечно, от природы той функции, разложения которой мы ищем.

Как мы видели выше, функция, разлагающаяся на данном отрезке в степенной ряд, должна иметь в каждой внутренней точке этого отрезка производные всех порядков. Так как, обратно, для любой функции, обладающей этим свойством, ряд Тэйлора (3) формально может быть написан в соседстве любой точки a этого отрезка, то может возникнуть соблазн предполагать, что это условие и достаточно для разложимости данной функции в степенной ряд. Однако это не так. Прежде всего, может случиться, что формально написанный для данной функции ряд Маклорена (2) окажется расходящимся для всех $x \neq 0$; значительно более интересным является, однако, то, что даже в случае, когда построенный для данной функции ряд Маклорена сходится, он может иметь своей суммой совсем не эту, а другую функцию; пример этого случая, ставший до известной степени классическим, дает функция

$$\varphi(x) = \begin{cases} e^{-x^{-2}} & (x \neq 0), \\ 0 & (x = 0). \end{cases}$$

Как легко вычислить, при $x = 0$ сама функция и ее производные всех порядков обращаются в нуль, вследствие чего все коэффициенты ряда Маклорена (а значит, и сумма его при всех значениях x) равны нулю.

Отсюда, между прочим, следует, что если данная функция, как мы видели, может разлагаться в степенной ряд не более чем одним способом, то, напротив, всякий сходящийся степенной ряд служит рядом Маклорена не для одной, а для бесчисленного множества функций. В самом деле, пусть $f(x)$ — сумма данного ряда, который служит, как мы видели, рядом Маклорена для этой функции; очевидно, что в таком случае все функции семейства $f(x) + \alpha\varphi(x)$, где α — любое вещественное число, а $\varphi(x)$ — функция, которую мы определили

выше, будут иметь данный ряд в качестве своего ряда Маклорена. Разумеется, *изображаться* данным рядом (т. е. быть его суммой) может только одна из функций этого семейства.

Ряды многочленов. Теорема Вейерштрасса. Так как последовательные частичные суммы степенного ряда представляют собой многочлены все возрастающих степеней, то всякая функция $f(x)$, разлагающаяся в степенной ряд, может быть приближенно представлена многочленом с любой степенью точности. Более точно это следует выразить так: если функция $f(x)$ изображается степенным рядом, радиус сходимости которого равен r, то для любого положительного $\rho < r$ и любого $\varepsilon > 0$ существует многочлен, отличающийся от $f(x)$ менее чем на ε для всех точек x отрезка $[-\rho, \rho]$ (мы не можем утверждать того же для отрезка $(-r, r)$, хотя бы даже открытого, так как на этом отрезке сходимость данного степенного ряда может быть неравномерной). Вспомним, что именно об этой задаче приближенного выражения функций через многочлены мы говорили в лекции V в связи с формулами Тэйлора и Маклорена.

Однако если всякая функция, разложимая в степенной ряд, может быть с любой степенью точности аппроксимирована с помощью многочлена, то обратное утверждение, по меньшей мере, не очевидно. Из того, что функция $f(x)$ на данном отрезке $[a, b]$ может быть представлена в виде многочлена с любой наперед заданной степенью точности, очевидно, еще не возникает разложимости ее где бы то ни было в степенной ряд. Это замечание имеет важное принципиальное значение; в самом деле, приближенное представление функций в виде многочленов является одним из важнейших орудий исследования этих функций и оперирования над ними; с другой стороны, мы знаем, что разложение в степенные ряды допускает лишь сравнительно узкий класс функций — в состав этого класса входит даже не всякая функция, имеющая производные всех порядков, что само по себе составляет уже весьма стеснительное требование; таким образом, если бы разложимость функции в степенной ряд была необходимым условием приближенной представимости ее с помощью многочле-

нов, то этой (весьма ценной) представимостью оказался бы обладающим лишь очень узкий класс функциональных зависимостей.

В действительности дело обстоит иначе. Оказывается, что необходимым и достаточным условием приближенной представимости данной фукции на некотором отрезке с любой степенью точности с помощью многочлена является *непрерывность* ее на этом отрезке. Таким образом, даже функции, нигде не дифференцируемые, могут допускать такое приближенное представление. Необходимость непрерывности, конечно, очевидна: функция, приближенно могущая быть представленной многочленом с любой степенью точности, есть предел равномерно сходящейся последовательности (и, значит, сумма равномерно сходящегося ряда) многочленов, т. е. непрерывных функций, и, следовательно, в силу известной теоремы (лекция IV, стр. 117) также должна быть непрерывной. То, что это необходимое условие вместе с тем является и достаточным, представляет собой один из самых глубоких и важных фактов математического анализа и составляет содержание знаменитой *теоремы Вейерштрасса*, к доказательству которой мы и переходим [1]). При этом, как вы сейчас увидите, нам придется начать несколько издалека.

Рассмотрим интеграл

$$I_n = \int_{-1}^{1} (1-u^2)^n\, du;$$

очевидно, при любом $n = 1, 2, \ldots$ это — положительное число; разобьем область интеграции на две части, а именно будем рассматривать отдельно интеграл

$$K_n = \int_{-n^{-1/3}}^{n^{-1/3}} (1-u^2)^n\, du,$$

распространенный на (малый при большом n) отрезок

---

[1]) В настоящее время известен целый ряд весьма различных доказательств этой теоремы; мы здесь выбираем одно из наиболее интересных в методологическом отношении.

$\left(-\frac{1}{\sqrt[3]{n}}, \frac{1}{\sqrt[3]{n}}\right)$, и дополнительный интеграл

$$L_n = \int\limits_{n^{-1/3} \leqslant |u| \leqslant 1} (1-u^2)^n \, du,$$

представляющий собой, в сущности, сумму двух интегралов, распространенных соответственно на отрезки $\left(-1, -\frac{1}{\sqrt[3]{n}}\right)$ и $\left(\frac{1}{\sqrt[3]{n}}, 1\right)$. Когда n безгранично возрастает, область интеграции у K_n стягивается к точке 0, область же интеграции у L_n, напротив, расширяется, стремясь охватить собой весь отрезок $[-1, 1]$; мы убедимся теперь, что, несмотря на это, при больших значениях n величина K_n исчерпывает собой почти всю величину I_n, а на долю интеграла L_n остается лишь ничтожно малая часть этой величины; объясняется это, конечно, тем, что при большом n подынтегральная функция $(1-u^2)^n$ лишь для очень малых значений $|u|$ имеет значения, сколько-нибудь заметные, будучи при сколько-нибудь больших значениях $|u|$ ничтожно малой; график этой функции схематически представлен на рис. 33.

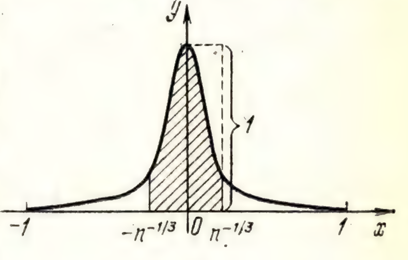


Рис. 33.

Лемма. *При* $n \to \infty$

$$\frac{K_n}{I_n} \to 1, \quad \frac{L_n}{I_n} \to 0.$$

Так как каждое из этих двух соотношений, очевидно, является следствием другого, то достаточно доказать какое-нибудь одно из них.

Подынтегральная функция в интеграле L_n имеет, очевидно, наибольшее значение при $|u| = n^{-1/3}$; так как область интеграции составляет часть отрезка $[-1, 1]$, то

$$L_n < (1 - n^{-2/3})^n \cdot 2.$$

С другой стороны, очевидно,

$$I_n > \int_{-\frac{1}{2}n^{-1/3}}^{\frac{1}{2}n^{-1/3}} (1-u^2)^n\,du > \left(1 - \frac{1}{4}n^{-2/3}\right)^n \cdot n^{-1/3}$$

и, следовательно,

$$\frac{L_n}{I_n} < 2n^{1/3}\left(\frac{1-n^{-2/3}}{1-\frac{1}{4}n^{-2/3}}\right)^n.$$

Но

$$\frac{1-n^{-2/3}}{1-\frac{1}{4}n^{-2/3}} = 1 - \frac{\frac{3}{4}n^{-2/3}}{1-\frac{1}{4}n^{-2/3}} <$$

$$< 1 - \frac{3}{4}n^{-2/3} < e^{-\frac{3}{4}n^{-2/3}}\ [1],$$

вследствие чего

$$\frac{L_n}{I_n} < 2n^{1/3}e^{-\frac{3}{4}n^{1/3}},$$

или, полагая $\frac{3}{4}n^{1/3} = z$,

$$\frac{L_n}{I_n} < \frac{8}{3}ze^{-z};$$

но при $n \to \infty$, конечно, $z \to +\infty$, а так как при этом $ze^{-z} \to 0$ [2], то

$$\lim_{n\to\infty} \frac{L_n}{I_n} = 0,$$

что и доказывает нашу лемму.

Какое же значение может иметь проведенное нами детальное исследование интеграла I_n для доказатель-

---

[1]) Известно, что $1 + x < e^x$ при любом $x \neq 0$; проще всего в этом убедиться, отыскивая наименьшее значение функции $e^x - 1 - x$.

[2]) Проще всего в этом убедиться, замечая, что при $z > 0$ $e^z = 1 + z + \frac{z^2}{2!} + \ldots > \frac{1}{2}z^2$, вследствие чего $ze^{-z} = \frac{z}{e^z} < \frac{2}{z}$.

ства столь общего предложения, каким является теорема Вейерштрасса? В математическом анализе это далеко не единичный пример такого случая, когда теорема чрезвычайно большой общности доказывается применением очень специального аналитического инструмента, каковым в нашем примере служит интеграл I_n; то специфическое свойство этого интеграла, которое делает его удобным орудием установления теоремы Вейерштрасса, находит себе выражение именно в доказанной нами лемме. Вы видите, таким образом, что излагаемое нами доказательство теоремы Вейерштрасса может служить методологически поучительным примером аналитического рассуждения.

Пусть теперь нам задана произвольная функция $f(x)$, непрерывная на отрезке $[0, 1]$. Интеграл

$$P_n(x) = \frac{1}{I_n}\int_0^1 f(v)\{1-(v-x)^2\}^n\,dv,$$

рассматриваемый как функция от x, представляет собой, очевидно, многочлен степени $2n$ (ибо таким многочленом является подынтегральная функция). Преобразуя переменную интеграции с помощью подстановки $v = x + u$, мы находим

$$P_n(x) = \frac{1}{I_n}\int_{-x}^{1-x} f(u+x)(1-u^2)^n\,du.$$

Пусть $0 < \alpha < \beta < 1$ и $\alpha \leqslant x \leqslant \beta$; тогда $-x \leqslant -\alpha < 0 < 1-\beta \leqslant 1-x$; поэтому, если n настолько велико, что $n^{-1/3}$ не превосходит меньшего из чисел α и $1-\beta$, то

$$-x < -n^{-1/3} < n^{-1/3} < 1-x \qquad (\alpha \leqslant x \leqslant \beta).$$

В последнем равенстве мы можем написать

$$\int_{-x}^{1-x} = \int_{-x}^{-n^{-1/3}} + \int_{-n^{-1/3}}^{n^{-1/3}} + \int_{n^{-1/3}}^{1-x}.$$

Но, обозначая через M наибольшее значение функции

$|f(x)|$ на отрезке $[0, 1]$, мы, очевидно, имеем

$$\frac{1}{I_n}\left|\left\{\int\limits_{-x}^{-n^{-1/3}} + \int\limits_{n^{-1/3}}^{1-x}\right\} f(u+x)(1-u^2)^n\,du\right| \leqslant$$

$$\leqslant \frac{M}{I_n}\left\{\int\limits_{-1}^{-n^{-1/3}} + \int\limits_{n^{-1/3}}^{1}\right\}(1-u^2)^n\,du = \frac{ML_n}{I_n} \to 0$$

при $n \to \infty$ в силу доказанной леммы; при этом, так как $\frac{ML_n}{I_n}$ от x не зависит, левая часть этого соотношения при $n \to \infty$ стремится к нулю *равномерно* относительно x на всем отрезке $[\alpha, \beta]$. Полагая

$$\frac{1}{I_n}\left\{\int\limits_{-x}^{-n^{-1/3}} + \int\limits_{n^{-1/3}}^{1-x}\right\} f(u+x)(1-u^2)^n\,du = R_n(x),$$

мы имеем, таким образом,

$$P_n(x) = \frac{1}{I_n}\int\limits_{-n^{-1/3}}^{n^{-1/3}} f(u+x)(1-u^2)^n\,du + R_n(x),$$

причем $R_n(x) \to 0$ при $n \to \infty$ равномерно на отрезке $[\alpha, \beta]$.

Вы уже, вероятно, догадались, что именно в лице многочлена $P_n(x)$ мы хотим найти многочлен, приближенно выражающий функцию $f(x)$; если так, то вам должен быть ясен и дальнейший план доказательства: ввиду равномерной малости величины $R_n(x)$ при больших значениях n надо только показать, что первое слагаемое правой части последнего равенства при этом на всем отрезке $[\alpha, \beta]$ близко к $f(x)$; это же почти очевидно, ибо $f(u+x)$ под интегралом мало отличается от $f(x)$, так как значения $|u|$, входящие в промежуток интеграции, ничтожно малы, заменяя же $f(u+x)$ через $f(x)$, мы тем самым заменяем все первое слагаемое выражением $f(x)\frac{K_n}{I_n}$, которое в силу нашей леммы стремится к $f(x)$ при $n \to \infty$. Чтобы завер-

шить доказательство, нам остается только провести это рассуждение со всей формальной строгостью; проще всего воспользоваться для этого первой теоремой о среднем значении (лекция VI, стр. 188), в силу которой

$$\int_{-n^{-1/3}}^{n^{-1/3}} f(u+x)(1-u^2)^n\, du = f(x+\theta n^{-1/3}) K_n,$$

где $-1 < \theta < 1$; мы находим, таким образом,

$$|f(x) - P_n(x)| = \left| f(x) - f(x+\theta n^{-1/3}) \frac{K_n}{I_n} - R_n(x) \right| \leqslant$$

$$\leqslant \left| f(x) - f(x+\theta n^{-1/3}) \frac{K_n}{I_n} \right| + |R_n(x)|.$$

Пусть $\varepsilon > 0$ произвольно мало; как мы уже знаем, при достаточно большом n

$$|R_n(x)| < \frac{\varepsilon}{2} \qquad (\alpha \leqslant x \leqslant \beta);$$

с другой стороны, так как $\frac{K_n}{I_n} \to 1$ при $n \to \infty$, а функция $f(x)$ равномерно непрерывна на всем отрезке $[0, 1]$, то при достаточно большом n

$$\left| f(x) - f(x+\theta n^{-1/3}) \frac{K_n}{I_n} \right| < \frac{\varepsilon}{2} \quad (\alpha \leqslant x \leqslant \beta);$$

таким образом, при достаточно большом n

$$|f(x) - P_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad (\alpha \leqslant x \leqslant \beta);$$

другими словами, при $n \to \infty$ многочлен $P_n(x)$ равномерно стремится к функции $f(x)$ на отрезке $[\alpha, \beta]$.

В основном этим теорема Вейерштрасса доказана; нам остается лишь избавиться от нескольких малосущественных ограничений; во-первых, мы должны перейти от специального отрезка $[0, 1]$ к произвольному отрезку $[a, b]$, и, во-вторых, мы должны показать, что требуемая теоремой Вейерштрасса равномерная аппроксимация имеет место не только на любом отрезке $[a', b']$, целиком лежащем внутри $[a, b]$ (именно так

у нас ведь обстояло дело, ибо наш отрезок $[\alpha, \beta]$ был любым отрезком, целиком лежащим внутри $[0, 1]$), но и на всем отрезке $[a, b]$.

Для достижения первой из этих двух целей положим

$$\frac{x-a}{b-a}=y, \quad x=a+(b-a)\,y \qquad (a \leqslant x \leqslant b),$$

$$f(x)=f[a+(b-a)\,y]=\varphi(y).$$

Когда x пробегает отрезок $[a, b]$, y пробегает отрезок $[0, 1]$, и если функция $f(x)$ непрерывна на отрезке $[a, b]$, то функция $\varphi(y)$ непрерывна на отрезке $[0, 1]$. Пусть $a < a' < b' < b$; положим

$$\frac{a'-a}{b-a}=\alpha, \quad \frac{b'-a}{b-a}=\beta, \qquad 0<\alpha<\beta<1.$$

В силу того, что мы только что доказали, для любого $\varepsilon > 0$ найдется такой многочлен $P_n(y)$, что

$$|\varphi(y)-P_n(y)|<\varepsilon \qquad (\alpha \leqslant y \leqslant \beta);$$

иначе,

$$\left|f(x)-P_n\left(\frac{x-a}{b-a}\right)\right|<\varepsilon \quad (a' \leqslant x \leqslant b');$$

но очевидно, что

$$\Pi_n(x)=P_n\left(\frac{x-a}{b-a}\right)$$

относительно x есть многочлен той же степени, что и $P_n(x)$; таким образом, функция $f(x)$, непрерывная на произвольном отрезке $[a, b]$, равномерно аппроксимируется многочленами с любой степенью точности на любом отрезке $[a', b']$, целиком лежащем внутри отрезка $[a, b]$.

Чтобы снять, наконец, и последнее ограничение, допустим снова, что функция $f(x)$ непрерывна на отрезке $[a, b]$, и положим

$$F(x)=\begin{cases} f(a) & (a-1 \leqslant x \leqslant a), \\ f(x) & (a \leqslant x \leqslant b), \\ f(b) & (b \leqslant x \leqslant b+1). \end{cases}$$

Очевидно, что функция $F(x)$ определена и непрерывна на отрезке $[a-1, b+1]$, внутри которого целиком ле-

жит отрезок $[a, b]$; в силу доказанного она равномерно аппроксимируется многочленами с любой степенью точности на отрезке $[a, b]$; но так как на этом отрезке $F(x) \equiv f(x)$, то же самое справедливо и для функции $f(x)$, чем теорема Вейерштрасса окончательно доказана.

Эту теорему можно формулировать и так: *всякая непрерывная функция есть сумма равномерно сходящегося ряда многочленов*; ибо, если, например, $P_n(x)$ означает (существующий в силу доказанного) многочлен, отличающийся от $f(x)$ менее чем на $\frac{1}{n}$ на всем протяжении отрезка $[a, b]$, то ряд

$$P_1(x) + \{P_2(x) - P_1(x)\} + \ldots + \{P_n(x) - P_{n-1}(x)\} + \ldots,$$

члены которого являются многочленами, очевидно, равномерно сходится к функции $f(x)$ на отрезке $[a, b]$.

Тригонометрические ряды. Тригонометрическим рядом называется ряд вида

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx). \qquad (4)$$

Так как все слагаемые этого ряда — периодические функции с периодом 2π, то такой же периодичностью обладает и сумма этого ряда. Поэтому, если функция, которую мы имеем в виду разложить в такой ряд, этой периодичностью не обладает, то имеет смысл искать ее разложения лишь на отрезках, длина которых не превосходит 2π. Вследствие этой же периодичности все рассмотрения, связанные с рядом (4), достаточно вести на каком-либо определенном отрезке длины 2π, например на отрезке $[-\pi, \pi]$.

Приведем прежде всего три соображения, в силу которых в некоторых случаях разложение функции в ряд типа (4) может оказаться более целесообразным, чем разложение в степенной ряд.

1) Как мы знаем, в степенной ряд разлагаются только функции, имеющие производные всех порядков (причем даже это стеснительное требование, вообще говоря, еще не оказывается достаточным); напротив, как мы увидим дальше, для разложимости функции в ряд вида (4) достаточно гораздо более скромных

предпосылок; в частности, такое разложение допускает, как мы увидим дальше, всякая функция, имеющая ограниченную и интегрируемую производную, причем и эта предпосылка не является необходимой.

2) Члены ряда (4) — периодические, волнообразные функции; некоторый элемент этой волнообразной периодичности сохраняют и частичные суммы этого ряда — черта, вообще говоря, отсутствующая у степенных рядов; поэтому, если исследуемая функция (как это часто бывает, например, в механике, физике, биологии, экономике) обнаруживает некоторую тенденцию к хотя бы приближенной волнообразности, то мы с полным основанием можем ожидать, что частичные суммы тригонометрического ряда будут имитировать поведение такой функции лучше, чем многочлены, каковыми являются частичные суммы степенных рядов.

3) Наконец, тригонометрические функции, представляющие собой члены ряда (4), обладают одним замечательным свойством, которое в чрезвычайной степени облегчает как исследование тригонометрических рядов, так и оперирование с ними и которого совершенно лишены члены степенных рядов. Это свойство есть так называемая *ортогональность* системы функций

$$1, \sin nx, \cos nx \qquad (n = 1, 2, \ldots); \tag{5}$$

оно состоит в том, что интеграл произведения двух любых различных между собой функций этой системы, взятый по любому отрезку длины 2π, равен нулю (свойство это вы либо уже знаете, либо легко докажете, преобразуя стоящее под знаком интеграла произведение в сумму на основании элементарных тригонометрических формул). Важность этого замечательного свойства трудно переоценить; даже гораздо более сложные функции, чем тригонометрические, если только они образуют ортогональную систему, во многих случаях оказываются в качестве элементов ряда исключительно удобным инструментом для исследования функций; многие свойства тригонометрической системы (5) оказываются общими свойствами любых ортогональных систем, что привело к созданию целой особой теории таких систем, в настоящее время уже широко развитой

и насчитывающей в своем составе много важных и глубоких аналитических фактов.

Коэффициенты Фурье. В случае степенных рядов мы видели, что коэффициенты ряда легко определяются, если известны значения разлагаемой функции и ее производных при $x = 0$. Этот же вопрос естественно поставить в первую очередь и для тригонометрических рядов. Пусть функция $f(x)$ изображается на отрезке $[-\pi, \pi]$ *равномерно* сходящимся рядом (4). Умножая обе части равенства

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

на $\cos kx$, где k — одно из чисел последовательности 0, 1, 2, ..., мы, очевидно, получаем представление функции $f(x) \cos kx$ в виде некоторого также равномерно сходящегося ряда; поэтому, беря интегралы обеих частей полученного равенства в пределах от $-\pi$ до π, мы можем интегрировать *почленно* ряд, стоящий в правой части (см. лекцию IV, стр. 120). Но в силу ортогональности системы тригонометрических функций интегралы всех членов этого ряда обратятся в нуль, за исключением единственного интеграла

$$\int_{-\pi}^{\pi} a_k \cos^2 kx \, dx = a_k \int_{-\pi}^{\pi} \frac{1 + \cos 2kx}{2} \, dx = \pi a_k; \qquad (6)$$

при этом мы предположили, что $k > 0$; при $k = 0$ мы получаем в качестве единственного отличного от нуля интеграла

$$\int_{-\pi}^{\pi} \frac{a_0}{2} \, dx = \pi a_0,$$

так что общий результат (6) сохраняется и в этом случае (именно с этой целью первый член ряда (4) и пишут обычно в виде $\frac{a_0}{2}$); таким образом, мы имеем при

любом $k \geqslant 0$

$$\int_{-\pi}^{\pi} f(x) \cos kx \, dx = \pi a_k,$$

или

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx \qquad (k = 0, 1, 2, \ldots); \qquad (7)$$

подобным же образом мы находим

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \, dx \qquad (k = 1, 2, \ldots). \qquad (8)$$

Формулы (7) и (8) и решают поставленную нами задачу — выражают коэффициенты ряда (4) через изображаемую этим рядом функцию. Эти формулы, как вы видите, сами по себе не требуют от функции $f(x)$ даже существования первой производной; коэффициенты a_k и b_k могут быть построены по этим формулам для любой интегрируемой функции; их называют обычно *коэффициентами Фурье* данной интегрируемой функции $f(x)$, а построенный с их помощью ряд (4) — *рядом Фурье* этой функции (хотя формулы (7) и (8) впервые были найдены Эйлером). Однако из этого, конечно, никак не следует (логическая ситуация здесь полностью аналогична той, с какой мы встретились при исследовании степенных рядов), что построенный для нашей интегрируемой функции ряд Фурье должен сходиться; а если он сходится, то из предыдущего никак не вытекает, что сумма его должна совпадать с функцией $f(x)$. И в действительности существует сколько угодно случаев, когда построенный для интегрируемой функции ряд Фурье оказывается расходящимся. Все, что установлено нами до сих пор, можно формулировать так: *если функция $f(x)$ может быть изображена равномерно сходящимся на отрезке $[-\pi, \pi]$ тригонометрическим рядом, то этот ряд есть ее ряд Фурье,* т. е. коэфициенты его выражаются через данную фукцию посредством формул (7) и (8). В случае ряда Маклорена мы пришли к совершенно аналогичному заключению, с той толь-

ко разницей, что там нам не приходилось особо требовать равномерной сходимости, так как из самой природы степенного ряда вытекает, как мы знаем, равномерная сходимость его на всяком отрезке, целиком помещающемся внутри интервала сходимости.

Очевидно, что основная задача теории тригонометрических рядов состоит в установлении тех предпосылок, при соблюдении которых данная функция является суммой своего ряда Фурье. Этой проблемой занималось очень большое число исследователей, ей посвящена огромная литература; в дальнейшем мы рассмотрим несколько простейших результатов, связанных с этим вопросом.

Аппроксимация в среднем. Но прежде покажем, как коэффициенты Фурье естественно появляются в связи с решением совсем другой задачи. Пусть мы имеем ограниченную и интегрируемую на отрезке $[-\pi, \pi]$ функцию $f(x)$; желая найти для нее приближенное выражение в виде «тригонометрического многочлена порядка n»

$$\Pi_n(x) = \frac{\alpha_0}{2} + \sum_{k=1}^{n} (\alpha_k \cos kx + \beta_k \sin kx),$$

спросим себя, как надлежит выбрать коэффициенты α_k, β_k, чтобы сделать эту аппроксимацию возможно точной. Конечно, в такой постановке наш вопрос еще не имеет определенности; мы ведь не определили, какой величиной мы будем измерять доброкачественность полученного приближения. Ясно, что выбор этой величины в широких пределах является произвольным; так, мы могли бы принять за погрешность нашего приближения верхнюю грань разности

$$|f(x) - \Pi_n(x)|$$

для $-\pi \leqslant x \leqslant \pi$; другой возможный выбор величины, оценивающей погрешность того или другого приближения, приписывает этой величине выражение

$$\int_{-\pi}^{\pi} |f(x) - \Pi_n(x)|\, dx;$$

наконец, можно оценивать эту погрешность и интегралом

$$\int_{-\pi}^{\pi} \{f(x) - \Pi(x)\}^2 \, dx; \tag{9}$$

последнее выражение с формальной стороны наиболее удобно как в теоретических изысканиях, так и в практических расчетах, ибо оно не пользуется символом абсолютного значения, очень часто приносящим затруднения при аналитических операциях. На этой последней форме оценки погрешности мы и остановимся; аппроксимацию, связанную с таким определением величины погрешности, в математике принято называть «аппроксимацией в среднем».

Теперь перед нами стоит совершенно определенная задача выбрать числа α_k, β_k таким образом, чтобы интеграл (9) получил наименьшее возможное значение. Очевидно, этот интеграл есть функция $2n+1$ переменных $\alpha_0, \alpha_1, \beta_1, \ldots, \alpha_n, \beta_n$, так что мы имеем дело с многомерной задачей на минимум; специальная природа этой задачи позволяет, однако, при решении ее обойтись без методов дифференциального исчисления.

Представим интеграл (9) в виде

$$\int_{-\pi}^{\pi} f^2(x)\, dx + \int_{-\pi}^{\pi} \Pi_n^2(x)\, dx - 2\int_{-\pi}^{\pi} f(x)\,\Pi_n(x)\, dx.$$

Обозначая через a_k, b_k коэффициенты Фурье функции $f(x)$, мы в силу формул (7) и (8), очевидно, имеем

$$\int_{-\pi}^{\pi} f(x)\,\Pi_n(x)\, dx = \pi\left\{\frac{\alpha_0 a_0}{2} + \sum_{k=1}^{n}(\alpha_k a_k + \beta_k b_k)\right\}. \tag{10}$$

С другой стороны, если мы учтем, что

$$\int_{-\pi}^{\pi} \cos^2 kx\, dx = \int_{-\pi}^{\pi} \sin^2 kx\, dx = \pi,$$

то ортогональность системы (5) легко дает нам

$$\int_{-\pi}^{\pi} \Pi_n^2(x)\, dx = \pi\left\{\frac{\alpha_0^2}{2} + \sum_{k=1}^{n}\left(\alpha_k^2 + \beta_k^2\right)\right\}. \tag{11}$$

Сопоставляя формулы (10) и (11) и пользуясь элементарными соотношениями

$$\alpha_k^2 - 2a_k\alpha_k = (\alpha_k - a_k)^2 - a_k^2,$$
$$\beta_k^2 - 2b_k\beta_k = (\beta_k - b_k)^2 - b_k^2,$$

мы получаем

$$\int_{-\pi}^{\pi} \Pi_n^2(x) - 2\int_{-\pi}^{\pi} f(x)\,\Pi_n(x)\,dx =$$
$$= \pi\left\{\frac{(\alpha_0 - a_0)^2}{2} + \sum_{k=1}^{n}[(\alpha_k - a_k)^2 + (\beta_k - b_k)^2]\right\} -$$
$$- \pi\left\{\frac{a_0^2}{2} + \sum_{k=1}^{n}\left(a_k^2 + b_k^2\right)\right\},$$

и, следовательно,

$$\int_{-\pi}^{\pi}[f(x) - \Pi_n(x)]^2\,dx = \int_{-\pi}^{\pi} f^2(x)\,dx -$$
$$- \pi\left\{\frac{a_0^2}{2} + \sum_{k=1}^{n}\left(a_k^2 + b_k^2\right)\right\} +$$
$$+ \pi\left\{\frac{(\alpha_0 - a_0)^2}{2} + \sum_{k=1}^{n}[(\alpha_k - a_k)^2 + (\beta_k - b_k)^2]\right\}.$$

В правой части только последнее слагаемое зависит от чисел α_k и β_k; мы должны поэтому найти минимум этого последнего слагаемого; но все члены его неотрицательны, и потому оно получает свое наименьшее значение (нуль), когда все эти члены обращаются в нуль, т. е. когда

$$\alpha_0 = a_0,\ \alpha_k = a_k,\ \beta_k = b_k \quad (k = 1, 2, \ldots, n);$$

таким образом, решение поставленной нами задачи гласит, что *из всех тригонометрических многочленов порядка n наилучшую аппроксимацию в среднем данной ограниченной и интегрируемой функции $f(x)$ дает многочлен, коэффициентами которого служат коэффициенты Фурье этой функции.* Обозначая этот тригонометри-

ческий многочлен через $P_n(x)$, мы имеем

$$\int_{-\pi}^{\pi} [f(x) - P_n(x)]^2\,dx =$$

$$= \int_{-\pi}^{\pi} f^2(x)\,dx - \pi\left\{\frac{a_0^2}{2} + \sum_{k=1}^{n}\left(a_k^2 + b_k^2\right)\right\}.$$

Это соотношение, между прочим, доказывает одно очень важное предложение: так как левая часть, очевидно, неотрицательна, то мы имеем при любом n

$$\frac{a_0^2}{2} + \sum_{k=1}^{n}\left(a_k^2 + b_k^2\right) \leqslant \frac{1}{\pi}\int_{-\pi}^{\pi} f^2(x)\,dx.$$

Так как правая часть этого неравенства от n не зависит, то сумма, стоящая в левой части, при возрастании n остается ограниченной; другими словами, *ряд*

$$\sum_{k=1}^{\infty}\left(a_k^2 + b_k^2\right)$$

сходится, какова бы ни была ограниченная интегрируемая функция $f(x)$. *Отсюда, в частности, следует, что коэффициенты Фурье* a_k, b_k *каждой такой функции стремятся к нулю при* $k \to \infty$.

Эти результаты, полученные совершенно элементарным путем, имеют очень большое значение для теории рядов Фурье.

Замкнутость системы тригонометрических функций. Как мы видели в теории степенных рядов, один и тот же ряд может служить рядом Маклорена для бесчисленного множества различных функций. Возможно ли аналогичное явление для рядов Фурье? Вопрос этот тесно связан с одним замечательным свойством системы (5) тригонометрических функций. Если ряд (4) служит рядом Фурье для двух непрерывных[1]) на отрезке $[-\pi, \pi]$ и различных между собой функций $f_1(x)$

[1]) Если не налагать никаких требований непрерывности, то поставленная задача решается тривиально, так как две функции, отличающиеся друг от друга только в какой-либо одной точке, очевидно, имеют все коэффициенты Фурье соответственно совпадающими между собой.

и $f_2(x)$, то это означает, что все соответственные коэффициенты Фурье этих двух функций совпадают между собой, а следовательно, разность

$$f(x) = f_1(x) - f_2(x)$$

этих двух функций, не равная тождественно нулю на отрезке $[-\pi, \pi]$, имеет все коэффициенты Фурье равными нулю; но это, очевидно, в силу (7) и (8) равносильно утверждению, что функция $f(x)$ *ортогональна* всем функциям системы (5); таким образом, в этом случае ортогональная система (5) была бы, как говорят, *незамкнутой*, т. е. к ней можно было бы добавить новую, не тождественно равную нулю функцию так, чтобы расширенная система оставалась ортогональной. Очевидно, возможно и обратное заключение: если система (5) незамкнута, то добавочная функция $f(x)$, ортогональная всем функциям системы (5), очевидно, тем самым имеет все коэффициенты Фурье равными нулю; но тогда все функции семейства

$$f_1(x) + \alpha f(x),$$

где α — любое вещественное число, имеют один и тот же ряд коэффициентов Фурье.

Покажем, что на самом деле система (5) замкнута; другими словами, всякая непрерывная функция, ортогональная всем функциям системы (5), должна тождественно обращаться в нуль.

Пусть $f(x)$ — такая функция. Очевидно, что тогда для любого тригонометрического многочлена $T(x)$ мы имеем

$$\int_{-\pi}^{\pi} f(x)\, T(x)\, dx = 0.$$

Допустим, что $f(x)$ не есть тождественный нуль, и пусть для определенности $f(x) > 0$ при $x = \alpha$; тогда, как мы знаем из теории непрерывных функций, можно найти такие положительные числа c и δ, что при $-\pi < \alpha - \delta \leqslant x \leqslant \alpha + \delta < \pi$ [1])

$$f(x) > c.$$

[1]) Очевидно, что в наших условиях α можно считать внутренней точкой отрезка $[-\pi, \pi]$.

Рассмотрим теперь выражение

$$T_n(x) = \left(\frac{1+\cos(x-\alpha)}{2}\right)^n;$$

производя возведение в степень по формуле бинома, мы находим для функции $T_n(x)$ выражение вида

$$T_n(x) = \sum_{r=0}^{n} c_r [\cos(x-\alpha)]^n;$$

но, как известно из тригонометрии, любая степень $\cos^n x$ может быть выражена как линейная комбинация функций

$$1, \cos x, \cos 2x, \ldots, \cos nx$$

с постоянными коэффициентами (доказывается очень просто методом полной индукции); таким образом, мы получаем

$$T_n(x) = \sum_{r=0}^{n} d_r \cos r(x-\alpha);$$

замечая же, наконец, что

$$\cos r(x-\alpha) = \cos r\alpha \cos rx + \sin r\alpha \sin rx,$$

мы, очевидно, приходим к выражению

$$T_n(x) = \frac{\alpha_0}{2} + \sum_{r=0}^{n} (\alpha_r \cos rx + \beta_r \sin rx),$$

где α_r, β_r — постоянные коэффициенты. Таким образом, функция $T_n(x)$ при всяком n есть тригонометрический многочлен, вследствие чего

$$\int_{-\pi}^{\pi} f(x)\, T_n(x)\, dx = 0 \qquad (n = 1, 2, \ldots). \qquad (12)$$

С другой стороны, представим себе в основных чертах поведение функции $T_n(x)$ при больших значениях n; так как функция $\frac{1+\cos(x-\alpha)}{2}$ на отрезке $[-\pi, \pi]$ заключена, очевидно, между 0 и 1, причем она равна единице только при $x = \alpha$, то при больших значениях

n величина $T_n(x)$ неотрицательна, равна единице при $x = \alpha$ и ничтожно мала для всех значений, сколько-нибудь удаленных от α; другими словами, течение ее имеет график того типа, какой изображен на рис. 33 (стр. 223), с той только разницей, что отрезок $[-1, 1]$ надо заменить отрезком $[\alpha - \pi, \alpha + \pi]$ и максимум вместо точки $x = 0$ поместить в точке $x = \alpha$.

Отсюда ясен план наших дальнейших рассуждений. Так как при больших значениях n функция $T_n(x)$ ничтожно мала вне отрезка $[\alpha - \delta, \alpha + \delta]$, то и соответствующие части интеграла (12) ничтожно малы, вследствие чего знак его определяется знаком интеграла

$$\int_{\alpha-\delta}^{\alpha+\delta} f(x)\, T_n(x)\, dx; \tag{13}$$

а так как под этим интегралом $T_n(x) > 0$ и $f(x) > c$, то надо только убедиться, что откинутые нами части интеграла (12) по абсолютному значению малы сравнительно с положительным интегралом (13), чтобы установить, что интеграл (12) не может равняться нулю, и тем самым прийти к противоречию. Проделаем необходимые для этого оценки. Положим

$$\int_{\alpha-\delta}^{\alpha+} f(x)\, T_n(x)\, dx = I_1, \qquad \left\{\int_{-\pi}^{\alpha-\delta} + \int_{\alpha+\delta}^{\pi}\right\} f(x)\, T_n(x)\, dx = I_2,$$

так что

$$0 = \int_{-\pi}^{\pi} f(x)\, T_n(x)\, dx = I_1 + I_2.$$

Принимая во внимание равенство $\frac{1+\cos(x-\alpha)}{2} = \cos^2\frac{x-\alpha}{2}$, имеем

$$I_1 \geqslant c \int_{\alpha-\delta}^{\alpha+\delta} \cos^{2n}\frac{x-\alpha}{2}\, dx =$$

$$= c \int_{-\delta}^{\delta} \cos^{2n}\frac{y}{2}\, dy = 2c \int_{0}^{\delta} \left(1 - \sin^2\frac{y}{2}\right)^n dy;$$

но при данной области интеграции

$$0 \leqslant \sin\frac{y}{2} < 1, \quad 0 < \cos\frac{y}{2} \leqslant 1,$$

вследствие чего

$$I_1 > 2c\int_0^{\delta}\left(1-\sin\frac{y}{2}\right)^n\cos\frac{y}{2}\,dy,$$

или, полагая $\sin\frac{y}{2} = z$, $\frac{1}{2}\cos\frac{y}{2}\,dy = dz$,

$$I_1 > 4c\int_0^{\sin\frac{\delta}{2}}(1-z)^n\,dz = \frac{4c}{n+1}\left\{1-\left(1-\sin\frac{\delta}{2}\right)^{n+1}\right\};$$

а так как при достаточно большом n выражение в фигурных скобках, очевидно, больше $\frac{1}{2}$, то

$$I_1 > \frac{2c}{n+1}. \tag{14}$$

С другой стороны, если мы обозначим через M наибольшее значение функции $f(x)$ на отрезке $[-\pi, \pi]$ и заметим, что в пределах тех двух интегралов, которые составляют собой I_2,

$$\cos^2\frac{x-\alpha}{2} \leqslant \cos^2\frac{\delta}{2}$$

и, следовательно,

$$T_n(x) \leqslant \cos^{2n}\frac{\delta}{2},$$

то мы получим

$$|I_2| \leqslant M\cos^{2n}\frac{\delta}{2}(\alpha-\delta+\pi+\pi-\alpha-\delta) <$$

$$< 2\pi M\cos^{2n}\frac{\delta}{2} = 2\pi M\rho^n, \tag{15}$$

где положено

$$\rho = \cos^2\frac{\delta}{2} < 1.$$

Так как $I_1 + I_2 = 0$, то $|I_1| = |I_2|$, откуда в силу (14) и (15)

$$2\pi M\rho^n > \frac{2c}{n+1}, \quad (n+1)\rho^{n+1} > \frac{c\rho}{\pi M};$$

но это и создает искомое противоречие, ибо $n\rho^n \to 0$ при $n \to \infty$ [1]) и потому не может оставаться при любых n больше постоянного положительного числа $\frac{c\rho}{\pi M}$.

Таким образом, мы доказали замкнутость ортогональной системы (5). Как мы уже видели, отсюда следует, что данный тригонометрический ряд (4) может служить рядом Фурье не более чем одной непрерывной функции. В частности, если этот ряд сходится равномерно, сумма его и есть единственная непрерывная функция, имеющая этот ряд своим рядом Фурье.

Сходимость рядов Фурье для функций, имеющих ограниченную интегрируемую производную. Покажем теперь, что всякая функция $f(x)$, периодическая с периодом 2π и имеющая ограниченную и интегрируемую производную, разлагается в равномерно сходящийся тригонометрический ряд (который, следовательно, есть ее ряд Фурье).

Условимся обозначать коэффициенты Фурье функции $f(x)$ через a_k, b_k, а коэффициенты Фурье функции $f'(x)$ — через a'_k, b'_k. Тогда интеграция по частям при $k > 0$ дает

$$\pi a_k = \int_{-\pi}^{\pi} f(x)\cos kx\,dx =$$

$$= \left[\frac{f(x)\sin kx}{k}\right]_{-\pi}^{\pi} - \frac{1}{k}\int_{-\pi}^{\pi} f'(x)\sin kx\,dx = -\frac{\pi b'_k}{k},$$

и подобным же образом $b_k = \frac{a'_k}{k}$. В дальнейшем нам

---

[1]) Полагая $n \ln\frac{1}{\rho} = x$, имеем

$$n\rho^n = \frac{1}{\ln\frac{1}{\rho}}\, xe^{-x} \to 0 \quad (x \to +\infty),$$

как мы видели в сноске на стр. 224.

удобно будет применить так называемое неравенство Коши — Шварца:

$$\left\{\sum_{k=1}^{n} u_k v_k\right\}^2 \leqslant \sum_{k=1}^{n} u_k^2 \sum_{k=1}^{n} v_k^2,$$

имеющее место для любых вещественных чисел u_k, v_k и любого n [1]).

Замечая, что

$$a_k \cos kx + b_k \sin kx = \frac{1}{k}(-b_k' \cos kx + a_k' \sin kx),$$

мы в силу неравенства

$$(\alpha \cos \varphi + \beta \sin \varphi)^2 \leqslant \alpha^2 + \beta^2$$

находим при $m > n$

$$\left|\sum_{k=n}^{m} (a_k \cos kx + b_k \sin kx)\right|^2 =$$

$$= \left|\sum_{k=n}^{m} \frac{1}{k}(-b_k' \cos kx + a_k' \sin kx)\right|^2 \leqslant$$

$$\leqslant \sum_{k=n}^{m} \frac{1}{k^2} \sum_{k=n}^{m} (-b_k' \cos kx + a_k' \sin kx)^2 \leqslant$$

$$\leqslant \sum_{k=n}^{m} \frac{1}{k^2} \cdot \sum_{k=n}^{m} (a_k'^2 + b_k'^2).$$

Но ряд $\sum_{k=1}^{\infty} \frac{1}{k^2}$ сходится. Точно так же и ряд

$$\sum_{k=1}^{\infty} (a_k'^2 + b_k'^2)$$

---

[1]) Доказательство: $\sum_{k=1}^{n} (u_k x + v_k)^2$, как функция от x, есть трехчлен второй степени, никогда не отрицательный; следовательно, дискриминант его

$$4\left(\left\{\sum_{k=1}^{n} u_k v_k\right\}^2 - \sum_{k=1}^{n} u_k^2 \sum_{k=1}^{n} v_k^2\right) \leqslant 0.$$

сходится (см. стр. 236), так как a'_k и b'_k — коэффициенты Фурье ограниченной интегрируемой функции $f'(x)$. Следовательно, в правой части последнего неравенства оба множителя стремятся к нулю при $n \to \infty$ и $m > n$ (критерий Коши!); а значит, и

$$\lim_{\substack{n\to\infty \\ m>n}} \sum_{k=n}^{m} (a_k \cos kx + b_k \sin kx) = 0,$$

и притом равномерно относительно x, ибо правая часть предыдущего неравенства от x не зависит. Но это и означает в силу критерия Коши, что ряд (4) равномерно сходится на отрезке $[-\pi, \pi]$; обозначая сумму его через $s(x)$, мы видим, что ряд (4) является рядом Фурье для каждой из функций $f(x)$ и $s(x)$; но в таком случае эти функции, будучи непрерывными, должны, как мы знаем, тождественно совпадать; этим наше утверждение доказано полностью.

Современная теория тригонометрических рядов доказывает сходимость рядов Фурье для значительно более широких классов функций, чем класс, рассмотренный нами; однако очень далеко такое расширение простираться не может, так как известно, что даже среди непрерывных функций имеются такие, для которых ряды Фурье сходятся не при всех значениях x. Здесь мы, конечно, не имеем возможности подробнее заняться этим кругом вопросов. Заметим только, что теория тригонометрических рядов, которой посвящена огромная, в том числе и учебная, литература, до настоящего времени сохраняет в своем составе большое число важных проблем, которые еще не удалось решить, и потому, естественно, привлекает к себе усилия многих исследователей.

Распространение на произвольные отрезки. До сих пор у нас шла речь о функциях, определенных на отрезке $[-\pi, \pi]$, и только на этом отрезке мы искали разложение данной функции в тригонометрический ряд; при этом мы, в сущности, негласно предполагали еще, что $f(\pi) = f(-\pi)$, ибо лишь при этом условии функция $f(x)$ может изображаться рядом вида (4) во всех точках закрытого отрезка $[-\pi, \pi]$. Теперь мы должны

посмотреть, как можно освободиться от этих ограничений, ибо ясно, что теория тригонометрических рядов может получить сколько-нибудь широкие применения лишь в том случае, если мы сумеем разлагать в тригонометрические ряды функции, определенные на любых отрезках и не стесненные никакими требованиями периодичности.

Прежде всего, очевидно, что (как мы, впрочем, указывали еще в самом начале) во всем предшествующем ничего не изменится, если вместо отрезка $[-\pi, \pi]$ мы положим в основу наших рассуждений любой отрезок $[a, a+2\pi]$ длины 2π, лишь бы мы имели $f(a+ +2\pi)=f(a)$. Таким образом, в нашем изложении предъявлялись требования лишь к длине рассматриваемого отрезка, а не к его положению; от функции же требовалось совпадение ее значений в концах данного отрезка.

Пусть теперь мы хотим разложить в тригонометрический ряд функцию $f(x)$, определенную на совершенно произвольном отрезке $[a, b]$ и не подчиненную никаким условиям периодичности; мы будем только предполагать, как и прежде, что функция $f(x)$ дифференцируема в каждой точке отрезка $[a, b]$ и что ее производная ограничена и интегрируема на этом отрезке.

Допустим сначала, что $b-a<2\pi$. Очевидно, что мы можем бесчисленным множеством способов определить на отрезке $[a, a+2\pi]$ функцию $f^*(x)$, которая имела бы на этом отрезке ограниченную интегрируемую производную, удовлетворяла требованию $f^*(a+2\pi)= =f^*(a)$ и совпадала с функцией $f(x)$ во всех точках отрезка $[a, b]$. В силу того, что мы доказали выше, функция $f^*(x)$ разлагается на отрезке $[a, a+2\pi]$ в равномерно сходящийся тригонометрический ряд. Очевидно, что на отрезке $[a, b]$ этот ряд изображает функцию $f(x)$ и, таким образом, полностью решает поставленную нами задачу.

Пусть теперь $b-a>2\pi$. Тогда мы прежде всего возьмем любое число $b'>b$ и определим на отрезке $[a, b']$ функцию $f^*(x)$ так, чтобы она имела на этом отрезке ограниченную и интегрируемую производную, удовлетворяла требованию $f^*(b')=f^*(a)$ и совпадала

с $f(x)$ во всех точках отрезка $[a, b]$. Если мы, сверх того, озаботимся тем, чтобы было $f^{*\prime}(b') = f^{*\prime}(a)$ (что, конечно, всегда возможно), и будем периодически продолжать функцию $f^*(x)$ за пределы отрезка $[a, b']$ в обе стороны, то мы получим периодическую функцию $f^*(x)$ периода $b' - a = 2l > 2\pi$, обладающую производной, ограниченной и интегрируемой на любом отрезке, и совпадающую с $f(x)$ на отрезке $[a, b]$.

Положим теперь

$$x = \frac{l}{\pi} y, \quad f^*(x) = f^*\left(\frac{l}{\pi} y\right) = \varphi(y).$$

Так как y пробегает отрезок $[-\pi, \pi]$, когда x изменяется от $-l$ до l, то функция $\varphi(y)$ обладает ограниченной интегрируемой производной на отрезке $[-\pi, \pi]$, причем $\varphi(-\pi) = f^*(-l) = f^*(l) = \varphi(\pi)$; в силу нашей теории ряд Фурье функции $\varphi(y)$ равномерно сходится к этой функции на отрезке $[-\pi, \pi]$:

$$\varphi(y) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos ny + b_n \sin ny) \quad (-\pi \leqslant y \leqslant \pi).$$

Но отсюда, очевидно $\left(y = \frac{\pi}{l} x\right)$,

$$f^*(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi}{l} x + b_n \sin\frac{n\pi}{l} x\right) \qquad (16)$$

равномерно на отрезке $-l \leqslant x \leqslant l$; так как функции $\cos\frac{n\pi}{l} x$ и $\sin\frac{n\pi}{l} x$ имеют, подобно функции $f^*(x)$, период $2l$, то разложение (16) имеет место равномерно на всей числовой прямой, откуда, в частности, на отрезке $[a, b]$ равномерно

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi}{l} x + b_n \sin\frac{n\pi}{l} x\right).$$

Итак, мы видим, что функция $f(x)$ изображается на отрезке $[a, b]$ тригонометрическим рядом, отличающимся от ряда (4) лишь тем, что элементами разложения вместо функций $\cos nx$, $\sin nx$ служат функции

$\cos n\frac{\pi}{l}x$, $\sin n\frac{\pi}{l}x$, имеющие период $2l$ вместо прежнего 2π. Этот результат решает поставленную нами задачу наилучшим возможным образом, ибо, конечно, заранее было ясно, что рядом вида (4), все члены которого имеют период 2π, функция $f(x)$ на отрезке $[a, b]$, вообще говоря, изображаться не может (в силу неравенства $b-a>2\pi$).

Мы должны еще посмотреть, как можно найти по функции $f(x)$ коэффициенты a_n, b_n изображающего ее ряда. Для этого заметим, что по первоначальному определению

$$a_k=\frac{1}{\pi}\int_{-\pi}^{\pi}\varphi(y)\cos ky\,dy,\quad b_k=\frac{1}{\pi}\int_{-\pi}^{\pi}\varphi(y)\sin ky\,dy.$$

Полагая $y=\frac{\pi}{l}x$, находим

$$a_k=\frac{1}{l}\int_{-l}^{l}\varphi\left(\frac{\pi}{l}x\right)\cos k\frac{\pi}{l}x\,dx=$$

$$=\frac{1}{l}\int_{-l}^{l}f^*(x)\cos k\frac{\pi}{l}x\,dx=\frac{1}{l}\int_{a}^{b'}f^*(x)\cos k\frac{\pi}{l}x\,dx,$$

и аналогично

$$b_k=\frac{1}{l}\int_{a}^{b'}f^*(x)\sin k\frac{\pi}{l}x\,dx.$$

Так как функций $f^*(x)$, удовлетворяющих всем поставленным нами требованиям, существует бесчисленное множество, то a_k и b_k заданием функции $f(x)$ на отрезке $[a, b]$ еще не определяются однозначно. Это не должно удивлять нас: ряд (16) призван изображать функцию $f(x)$ лишь в интервале $[a, b]$, длина которого меньше, чем период $2l$ элементов разложения; совершенно аналогично и в первоначально разобранном нами случае мы могли бы получить бесчисленное множество различных рядов вида (4), изображающих функцию $f(x)$ на любой части отрезка $[-\pi, \pi]$.

ЛЕКЦИЯ VIII
ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Основные понятия.— Существование решения.— Единственность решения.— Зависимость решений от параметров.— Преобразование переменных.— Системы уравнений и уравнения высших порядков.

Основные понятия. В конце нашей шестой лекции мы говорили о том, что в идейном плане целью всякого процесса интеграции является установление *суммарных «интегральных»* черт изучаемого явления по данным его *локальным* характеристикам. Задачи, идущие в этом направлении, весьма многочисленны во всех прикладных науках, пользующихся средствами математического анализа. Однако те весьма разнообразные запросы, которые предъявляют этого рода задачи к математическому аппарату, далеко не могут быть удовлетворены таким элементарным с идейной точки зрения орудием, каким является интеграл — обыкновенный или многомерный; этот простейший интегрирующий аппарат лишь в сравнительно немногих, наиболее примитивных, случаях оказываются достаточным для исчерпывающего решения поставленной задачи.

Чаще всего дело обстоит так, что известные нам локальные черты изучаемого явления приводят к построению *дифференциальных уравнений* (т. е. уравнений, связывающих между собой значения независимых переменных, функций и их производных различных порядков), в которых в качестве неизвестных фигурируют как раз функции, описывающие искомые интегральные черты этого явления; решение поставленной конкретной задачи математически сводится, таким образом, к решению некоторой системы дифференциаль-

ных уравнений, т. е. к определению входящих в эти уравнения неизвестных функций; эта математическая задача значительно более сложна, чем простое интегрирование функций, которое с точки зрения теории дифференциальных уравнений является лишь простейшим и притом тривиальным частным случаем решения дифференциальных уравнений; это находит себе выражение в том, что всякий раз, когда решение какого-либо типа дифференциальных уравнений тем или другим способом удается свести к простому интегрированию функций (или, как принято говорить в теории дифференциальных уравнений, к «квадратурам»), общая задача тем самым считается решенной.

Нам будет полезно рассмотреть на каком-нибудь очень простом примере, каким образом наше знакомство с локальными чертами явления может породить дифференциальное уравнение, решение которого затем позволяет дать этому явлению и суммарное интегральное описание.

Представьте себе сосуд вместимостью a *л*, наполненный водой, в которой растворена какая-либо соль; пусть через этот сосуд непрерывно протекает жидкость, причем в него вливается в единицу времени b *л* чистой воды и из него вытекает столько же литров раствора; пусть еще известно, что в некоторый начальный момент времени $t = 0$ сосуд содержал c *кг* соли; мы хотим узнать, какое число x *кг* соли наш сосуд будет содержать по истечении t единиц времени [1]).

Какие локальные черты явления нам даны в поставленной задаче? Мы знаем, что раствор вытекает из сосуда со скоростью b *л* в единицу времени; в данный момент времени t сосуд содержит неизвестное нам число x *кг* соли, а так как вместимость сосуда составляет a *л*, то каждый литр раствора содержит $\frac{x}{a}$ *кг* соли, а b *л* содержат $\frac{bx}{a}$ *кг*. Значит, если бы в течение единицы времени, начиная с момента t, концентрация раствора оста-

[1]) При этом мы предполагаем перемешивание раствора столь быстрым, что практически можно считать концентрацию соли одной и той же во всех точках сосуда в каждый данный момент.

валась постоянной (такой, какова она в момент t), то количество соли в сосуде за эту единицу времени уменьшилось бы на $\frac{bx}{a}$ *кг*; такова, следовательно скорость $-\frac{dx}{dt}$ уменьшения количества соли в сосуде для момента t (знак минус здесь необходим потому, что $\frac{dx}{dt}$ есть скорость *прироста* количества соли). Мы имеем

$$\frac{dx}{dt} = -\frac{bx}{a}; \qquad (1)$$

это соотношение, выражающее мгновенную скорость прироста количества соли $\frac{dx}{dt}$ через наличное в данный момент ее количество x, и дает нам локальную (здесь, конечно, более удобно говорить: «моментальную» или «мгновенную») характеристику явления. Функция $x = x(t)$, которую мы хотим найти, есть неизвестная, искомая функция полученного уравнения.

Может ли эта функция быть найдена непосредственно методами интегрального исчисления? Нам дано выражение ее производной, а отыскивать функцию, когда дана ее производная, и есть ведь основное дело интегрального исчисления. Однако задача все же не вполне обычна: производная искомой функции выражена не через независимую переменную (как мы привыкли), а через саму неизвестную функцию; непосредственно интегральное исчисление решением подобного рода задач не занимается, и потому формально мы стоим перед принципиально новой задачей — решением *дифференциального уравнения* (1); однако в данном случае вопрос, конечно, тривиальным образом приводится к задаче интегрального исчисления; написав уравнение (1) в виде

$$\frac{dx}{x} = -\frac{b}{a}\,dt,$$

мы, как говорят, «разделяем» переменные; простая интеграция даст слева $\ln x$, справа $-\frac{b}{a}t$, так что

$$\ln x = -\frac{b}{a}t + k,$$

где k — постоянное число; для определения этого числа мы (это весьма характерно для всей практики решения дифференциальных уравнений) обращаемся к «начальному» условию: $x = c$ при $t = 0$; это дает нам $k = \ln c$, так что окончательно

$$x = ce^{-\frac{b}{a}t};$$

этим поставленная нами задача полностью решена; мы видим, что количество соли в сосуде убывает с течением времени по «показательному» закону.

Теперь представьте себе, что жидкость, вытекающая из нашего сосуда, протекает еще через второй сосуд той же вместимости, первоначально (т. е. в момент $t = 0$) заполненный чистой водой, и что здесь также втекает и вытекает по b литров жидкости в единицу времени. Очевидно, что в этот второй сосуд постоянно вводится тогда соль; требуется узнать число y *кг* соли, которое будет содержаться во втором сосуде в момент t.

Здесь опять нам первоначально даны только локальные (мгновенные) характеристические черты явления. Очевидно, в единицу времени во второй сосуд втекает столько же соли, сколько ее вытекает из первого, т. е., как мы знаем,

$$\frac{b}{a}x = \frac{bc}{a}e^{-\frac{b}{a}t} \text{ кг};$$

с другой стороны, в момент t каждый литр жидкости во втором сосуде содержит $\frac{y}{a}$ *кг* соли, а значит, те b литров, которые из него вытекают в единицу времени, содержат $\frac{b}{a}y$ *кг* соли; следовательно, общий прирост количества соли во втором сосуде в единицу времени для момента t равен

$$\frac{dy}{dt} = \frac{bc}{a}e^{-\frac{b}{a}t} - \frac{b}{a}y = \frac{b}{a}\left(ce^{-\frac{b}{a}t} - y\right). \qquad (2)$$

Вы видите, что и здесь локальное описание явления нашло свое математическое выражение в некотором дифференциальном уравнении (2); снова нам дает-

ся выражение производной $\frac{dy}{dt}$ искомой функции y; однако на этот раз решение полученного уравнения уже далеко не так легко найти, как прежде; уравнение (2) дает нам выражение производной $\frac{dy}{dt}$, содержащее как независимую переменную t, так и искомую функцию y, и «разделить» переменные, как мы это сделали в уравнении (1), здесь непосредственно не удается. Таким образом, мы стоим перед принципиально новой задачей, которая, правда, в данном конкретном случае может быть решена сравнительно просто, но для решения которой в общем случае у нас никаких методов нет.

Как же выглядит этот «общий случай», как определяется общее понятие дифференциального уравнения? В наших примерах мы имели дело с уравнениями вида

$$\frac{dy}{dx} = f(x, y), \tag{3}$$

где x — независимая переменная, а y — неизвестная функция от x; функция $f(x, y)$, конечно, задана нам. Задача состоит в отыскании функции $y = \varphi(x)$, удовлетворяющей уравнению (3), т. е. такой, что тождественно на некотором отрезке $a \leqslant x \leqslant b$

$$\varphi'(x) = f(x, \varphi(x)).$$

Несколько более общий тип дифференциального уравнения представляет собой соотношение

$$F(x, y, y') = 0,$$

где снова x — независимая переменная, y — неизвестная функция от x и $y' = \frac{dy}{dx}$. Как и прежде, разыскивается функция $y = \varphi(x)$, тождественно удовлетворяющая соотношению

$$F(x, \varphi(x), \varphi'(x)) = 0.$$

Далее, бывают случаи, когда локальные черты изучаемого явления требуют для своего выражения не только первой производной искомой функции, но и ее производных высших порядков, так что дифференциаль-

ное уравнение задачи получает вид

$$F(x, y, y', \ldots, y^{(n)}) = 0; \qquad (4)$$

такое уравнение называется дифференциальным уравнением *порядка* n; очевидно, это есть самый общий тип дифференциального уравнения для задач, имеющих дело лишь с одной неизвестной функцией и одной независимой переменной. Однако таково положение вещей лишь в простейших случаях; часто приходится иметь дело с *несколькими* искомыми функциями, зависящими от *нескольких* независимых переменных.

Пусть сначала речь идет о любом числе неизвестных функций $y_1, y_2, \ldots, y_k$, зависящих, однако, лишь от одной независимой переменной x; чтобы задача получила определенность, локальные характеристики явления должны приводить нас к *системе* дифференциальных уравнений, число которых равно числу k неизвестных функций. Общий вид такой системы уравнений, очевидно, таков:

$$F_i\left(x, y_1, y_1', \ldots, y_1^{(n_1)}, y_2, y_2', \ldots, y_2^{(n_2)}, \ldots \right.$$
$$\left. \ldots, y_k, y_k', \ldots, y_k^{(n_k)}\right) = 0 \qquad (i = 1, 2, \ldots, k).$$

Вместе с тем мы получили и наиболее общую задачу теории так называемых «обыкновенных» дифференциальных уравнений; этим термином обычно принято характеризовать дифференциальные уравнения и системы уравнений, содержащие лишь одну независимую переменную.

Хорошим примером может служить (известная вам, конечно) система уравнений движения материальной точки

$$m\frac{d^2x}{dt^2} = X\left(t, x, y, z, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right),$$

$$m\frac{d^2y}{dt^2} = Y\left(t, x, y, z, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right),$$

$$m\frac{d^2z}{dt^2} = Z\left(t, x, y, z, \frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right),$$

где единственной независимой переменной служит время t, а неизвестными функциями являются коор-

динаты x, y, z материальной точки; m здесь означает массу этой материальной точки, а X, Y, Z — компоненты равнодействующей приложенных к ней сил, в общем случае зависящие от времени, положения и скорости движущейся точки; начальными данными могут служить три координаты и три компоненты скорости этой точки в некоторый определенный «начальный» момент времени $t = t_0$.

Если независимых переменных имеется несколько, то мы имеем дело с *уравнением* (или системой уравнений) *с частными производными*, неизвестные функции являются теперь функциями нескольких переменных, и уравнения, естественно, содержат их частные производные по этим переменным. Теория уравнений с частными производными, как легко понять, значительно сложнее теории «обыкновенных» уравнений; здесь мы совсем не будем ее касаться.

Но и теория обыкновенных дифференциальных уравнений не богата сколько-нибудь эффективными общими методами, которые позволили бы находить решения более или менее широких классов таких уравнений; отчасти это легко и предвидеть, так как ведь даже обычное интегрирование, примененное к элементарным функциям, во многих случаях приводит, как мы знаем, к функциям неэлементарным; тем более этого следует ожидать от задачи более общей и сложной. Однако, если мы даже станем на ту обычную для теории дифференциальных уравнений точку зрения, что, сведя решение задачи к простым интеграциям, мы можем считать эту задачу решенной,— даже принимая эту точку зрения, мы получаем возможность продвинуться вперед лишь очень недалеко, ибо такая редукция к «квадратурам» удается лишь для немногих простейших (правда, с чисто практической точки зрения наиболее важных) типов дифференциальных уравнений. Мы здесь не будем не только исследовать, но даже и перечислять эти типы,— по этому вопросу вы найдете сколько угодно материала в любом, даже самом элементарном, учебном руководстве. Нас же, естественно, должны занять проблемы более принципиального характера.

Существование решения. Мы уже говорили о том, что основная задача интегрального исчисления может рассматриваться как простейший частный случай решения дифференциального уравнения; именно, если в уравнении

$$\frac{dy}{dx} = f(x, y) \tag{3}$$

правая часть не зависит от y, мы получаем уравнение вида

$$\frac{dy}{dx} = f(x), \tag{3'}$$

решение которого, очевидно, эквивалентно интегрированию функции $f(x)$. По поводу этой задачи мы имели случай убедиться в том, что, даже если функция $f(x)$ непрерывна, существование интеграла (т. е. решения уравнения (3')) доказывается специальным рассуждением (опирающимся на равномерную непрерывность функции $f(x)$); если же функция $f(x)$ ограничена, но имеет точки разрыва, то интеграла, вообще говоря, не существует.

Из всего этого заранее ясно, что вопрос о существовании решений уравнения (3) должен быть тем более в достаточной мере сложным и во всяком случае требует специального исследования; в еще большей степени это относится, конечно, к тем более общим типам уравнений и систем уравнений, которые мы приводили выше. Чтобы идейная сторона рассмотрения этой основной для теории дифференциальных уравнений проблемы могла выступить перед вами достаточно рельефно, мы должны по возможности разгрузить наше изложение от сложностей чисто технической природы; поэтому мы в дальнейшем ограничимся рассмотрением уравнений типа (3), т. е. уравнений первого порядка, разрешенных относительно производной неизвестной функции.

Пусть функция $f(x, y)$ непрерывна в некоторой области D плоскости xy; мы покажем, что в таком случае для каждой точки (x_0, y_0) этой области существует функция $y = \varphi(x)$, удовлетворяющая уравнению (3)

в некоторой окрестности точки x_0 *и притом такая, что* $y_0 = \varphi(x_0)$. *Очевидно, область* D *можно считать замкнутой и ограниченной.*

Для доказательства нам потребуется одно вспомогательное предложение, представляющее, впрочем, и самостоятельный интерес. Пусть мы имеем любую последовательность функций $S = \{F_1(x), F_2(x), \ldots \ldots, F_\nu(x), \ldots\}$, определенных на некотором отрезке $[a, b]$. Условимся называть последовательность S *ограниченной* на этом отрезке, если существует такое положительное число M, что $|F_\nu(x)| < M$ для любого x отрезка $[a, b]$ и для любого ν; будем, далее, называть последовательность S *равностепенно непрерывной* на отрезке $[a, b]$, если для любого $\varepsilon > 0$ найдется такое положительное число δ, что для любой функции $F_\nu(x)$ и для любой пары точек x_1, x_2 отрезка $[a, b]$, удовлетворяющих неравенству $|x_1 - x_2| < \delta$, имеет место неравенство $|F_\nu(x_1) - F_\nu(x_2)| < \varepsilon$. Очевидно, что если последовательность S ограничена (равностепенно непрерывна) на отрезке $[a, b]$, то и каждая из входящих в нее функций ограничена (непрерывна) на этом отрезке; обратное, конечно, вообще говоря, неверно: ограниченность и равностепенная непрерывность последовательности S требуют, сверх наличия соответствующего свойства у каждой функции этой последовательности, еще и *равномерности* этого свойства относительно всей совокупности функций данной последовательности.

Условимся, наконец, называть *шириной* последовательности S на отрезке $[a, b]$ верхнюю грань всех величин $|F_\nu(x) - F_\mu(x)|$, где x — любая точка отрезка $[a, b]$, а F_ν и F_μ — две произвольные функции последовательности S.

Лемма. *Всякая ограниченная и равностепенно непрерывная на отрезке* $[a, b]$ *последовательность функций* S *содержит равномерно сходящуюся на этом отрезке подпоследовательность.*

Доказательство проведем несколькими последовательными этапами.

1) Докажем прежде всего, что, как бы мало ни было положительное число ε, существует такое другое положительное число δ, что для любого отрезка $[\alpha, \beta] \subset$

$\subset [a, b]$, длина которого $< \delta$, последовательность S содержит подпоследовательность S', ширина которой на отрезке $[\alpha, \beta]$ меньше чем ε. В силу равностепенной непрерывности последовательности S на отрезке $[a, b]$ для любого натурального числа n найдется такое положительное число δ, что колебание любой функции последовательности S на любом отрезке длины $< \delta$ будет меньше чем $\frac{M}{n}$; отсюда следует, что все значения любой из этих функций на отрезке $[\alpha, \beta]$ будут помещаться в некотором отрезке длины $\frac{M}{n}$; пусть, на один момент, ξ, η $\left(\eta - \xi = \frac{M}{n}\right)$ означают концы одного из таких отрезков; очевидно, мы можем определить целое число k так, что $\frac{k-1}{n} M \leqslant \xi < \frac{k}{n} M$; в таком случае

$$\frac{k-1}{n} M \leqslant \xi < \eta < \frac{k+1}{n} M;$$

и так как отрезок $[\xi, \eta]$ принадлежит отрезку $[-M, M]$, то $-(n-1) \leqslant k \leqslant n-1$; мы видим, таким образом, что все значения, принимаемые любой из функций последовательности S на отрезке $[\alpha, \beta]$, принадлежат некоторому отрезку вида

$$\left[\frac{k-1}{n} M, \frac{k+1}{n} M\right] \quad (-(n-1) \leqslant k \leqslant n-1);$$

так как этих отрезков конечное число, а последовательность S содержит бесконечное число функций, то по меньшей мере один из этих отрезков будет содержать все значения некоторой подпоследовательности S' функций последовательности S; очевидно, что ширина последовательности S' на отрезке $[\alpha, \beta]$ не превосходит $\frac{2M}{n}$ и будет меньше чем ε, если мы выберем число n так, чтобы $\frac{2M}{n} < \varepsilon$.

2) Теперь убедимся, что последовательность S содержит и такую подпоследовательность S', ширина которой на всем отрезке $[a, b]$ меньше чем ε. С этой целью выберем натуральное число m столь большим,

чтобы было $\frac{b-a}{m}<\delta$; тогда в силу доказанного в пункте 1) последовательность S содержит подпоследовательность S_1, ширина которой меньше чем ε на отрезке $\left[a, a+\frac{b-a}{m}\right]$; на том же основании S_1 содержит подпоследовательность S_2, ширина которой меньше чем ε на отрезке $\left[a+\frac{b-a}{m}, a+2\frac{b-a}{m}\right]$, и т. д. В конечном счете мы приходим к подпоследовательности $S_m=S'$ последовательности S, ширина которой меньше чем ε на отрезке $\left[a+(m-1)\frac{b-a}{m}, b\right]$, а также и на каждом из предшествующих отрезков, так как $S_m\subset S_{m-1}\subset\ldots\subset S_2\subset S_1$; но это значит, что ширина последовательности S' меньше чем ε на всем отрезке $[a, b]$, т. е. наше утверждение доказано.

3) Теперь доказательство леммы завершается уже совсем просто. Обозначим через S_1 подпоследовательность последовательности S, ширина которой на $[a, b]$ меньше чем 1; через S_2 — подпоследовательность последовательности S_1, ширина которой на $[a, b]$ меньше чем $\frac{1}{2}$, и вообще через S_n — такую подпоследовательность последовательности S_{n-1}, ширина которой на $[a, b]$ меньше чем $\frac{1}{n}$ (все эти подпоследовательности существуют в силу доказанного в пункте 2)). Пусть $F_{\nu_1}(x)$ — любая функция из S_1, $F_{\nu_2}(x)$ — любая функция из S_2, вообще, $F_{\nu_n}(x)$ — любая функция из S_n, где $\nu_n>\nu_{n-1}$. Так как $S_{n+p}\subset S_n$ при любом $p>0$, то $F_{\nu_n}(x)\in S_n$ и $F_{\nu_{n+p}}(x)\in S_n$, вследствие чего для любой точки x отрезка $[a, b]$

$$|F_{\nu_n}(x)-F_{\nu_{n+p}}(x)|<\frac{1}{n}\qquad(p=0, 1, 2, \ldots).$$

Но в силу критерия Коши отсюда следует, что последовательность функций $F_{\nu_1}(x), F_{\nu_2}(x), \ldots, F_{\nu_n}(x), \ldots$ сходится равномерно на отрезке $[a, b]$, что и завершает доказательство леммы.

Теперь мы можем перейти к доказательству основной теоремы, сформулированной выше.

Но так как это доказательство будет основано на идее, которую проще всего усвоить в ее геометрическом истолковании, то нам будет очень полезно, прежде чем к нему приступить, посмотреть, какой вид получает в геометрической интерпретации самая задача отыскания решений данного уравнения (3).

Графические изображения решений $y = \varphi(x)$ этого уравнения обычно называют его *интегральными кривыми*. Уравнение (3) относит каждой точке (x, y) области D некоторое определенное направление, характеризуемое угловым коэффициентом $\frac{dy}{dx} = f(x, y)$; совокупность точек области D, с отнесенным каждой из них определенным направлением, образует так называемое *поле направлений*; именно это поле и задается уравнением (3). Интегральная кривая этого уравнения есть такая кривая, направление касательной в каждой точке которой совпадает с направлением поля в этой точке. Задача отыскания такого решения $y = \varphi(x)$ уравнения (3), которое обращается в y_0 при $x = x_0$, с геометрической точки зрения есть, таким образом, задача отыскания интегральной кривой данного уравнения, проходящей через точку (x_0, y_0); доказать существование этого решения означает поэтому показать, что через эту точку проходит по меньшей мере одна интегральная кривая.

Метод, которым мы сейчас воспользуемся для этой цели, будет состоять, как вы увидите, в построении вспомогательных кривых, по своим свойствам все более и более приближающихся к интегральной кривой, которую нам удастся получить из них с помощью предельного перехода.

Пусть (x_0, y_0) — внутренняя точка замкнутой ограниченной области D, в которой функция $f(x, y)$ непрерывна, и пусть M — верхняя грань функции $|f(x, y)|$ в этой области. Если число $\alpha > 0$ достаточно мало, то при $x_0 - \alpha \leqslant x \leqslant x_0 + \alpha$, $y_0 - M\alpha \leqslant y \leqslant y_0 + M\alpha$ точка (x, y) также будет принадлежать области D; мы покажем, что существует функция $y = \varphi(x)$, которая для

всех x отрезка $[x_0-\alpha,\ x_0+\alpha]$ удовлетворяет уравнению (3), причем $\varphi(x_0)=y_0$. Доказательство мы снова для более легкой обозримости разобьем на отдельные этапы.

1) Положим $x_k = x_0 + \frac{k}{n}\alpha$ $(0 \leqslant k \leqslant n)$, так что точки $x_1, x_2, \ldots, x_{n-1}$ делят отрезок $[x_0,\ x_0+\alpha]$ на n равных частей.

Построим теперь (рис. 34) на участке $[x_0,\ x_0+\alpha]$ ломаную линию $y=\varphi_n(x)$ с вершинами, имеющими

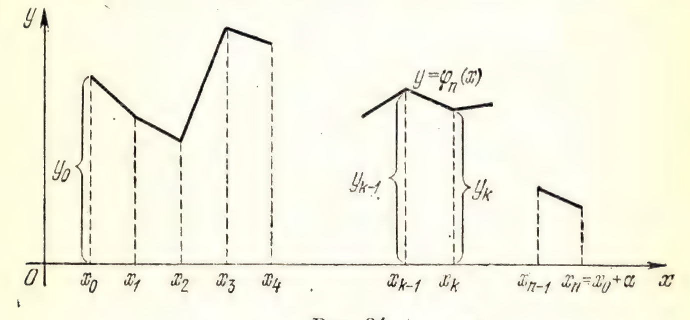


Рис. 34.

абсциссы $x_0, x_1, \ldots, x_n$, так, чтобы угловой коэффициент каждого звена совпадал с направлением поля в его левом конце. Очевидно, это можно сделать рекуррентно, переходом от x_{k-1} к x_k, причем ордината $y_k=\varphi_n(x_k)$ будет определяться рекуррентной формулой

$$y_k = y_{k-1} + f\left(x_{k-1}, y_{k-1}\right)\left(x_k - x_{k-1}\right) =$$
$$= y_{k-1} + f\left(x_{k-1}, y_{k-1}\right)\frac{\alpha}{n} \quad (1 \leqslant k \leqslant n), \qquad (5)$$

а уравнение ломаной на участке $x_{k-1} \leqslant x \leqslant x_k$ задается формулой

$$\varphi_n(x) = y_{k-1} + f(x_{k-1},\ y_{k-1})(x - x_{k-1}) \qquad (1 \leqslant k \leqslant n).$$

Чтобы убедиться, что ломаная $y=\varphi_n(x)$ на всем протяжении от x_0 до $x_n = x_0+\alpha$ принадлежит области D, достаточно, очевидно, показать, что

$$|y_k - y_0| < M\alpha \qquad (1 \leqslant k \leqslant n);$$

нам будет удобнее установить более сильное неравенство

$$|y_k - y_0| \leqslant \frac{k}{n} M\alpha \qquad (1 \leqslant k \leqslant n);$$

очевидно, оно выполнено для $k = 0$; если оно верно для некоторого $k < n$, то точка (x_k, y_k) принадлежит области D, и потому $|f(x_k, y_k)| \leqslant M$; но в таком случае рекуррентная формула (5) дает

$$|y_{k+1} - y_0| \leqslant |y_k - y_0| + \frac{\alpha}{n} |f(x_k, y_k)| \leqslant \frac{k+1}{n} M\alpha, \quad (6)$$

так что наше неравенство справедливо и для числа $k+1$. Таким образом,

$$|\varphi_n(x) - y_0| \leqslant M\alpha \qquad (x_0 \leqslant x \leqslant x_0 + \alpha); \quad (7)$$

вся ломаная $y = \varphi_n(x)$ лежит в области D; так как это имеет место для любого n, то, в частности, отсюда следует, что последовательность функций $\varphi_n(x)$ $(n = 1, 2, \ldots)$ ограничена на отрезке $[x_0, x_0 + \alpha]$.

2) Пусть x' и x'' — две точки отрезка $[x_0, x_0 + \alpha]$, и пусть для определенности

$$x_{k-1} \leqslant x' < x_k < \ldots < x_{l-1} \leqslant x'' < x_l;$$

тогда

$$\begin{aligned} \varphi_n(x'') - \varphi_n(x') = [\varphi_n(x'') - \varphi_n(x_{l-1})] + \\ + [\varphi_n(x_{l-1}) - \varphi_n(x_{l-2})] + \ldots + [\varphi_n(x_k) - \varphi_n(x')] = \\ = (x'' - y_{l-1}) f(x_{l-1}, y_{l-1}) + \\ + \frac{\alpha}{n} \{f(x_{l-2}, y_{l-2}) + \ldots + f(x_k, y_k)\} + \\ + (x_k - x') f(x_{k-1}, y_{k-1}). \quad (8) \end{aligned}$$

Из этого основного соотношения мы получим теперь ряд выводов.

Убедимся прежде всего, что последовательность функций $\varphi_n(x)$ $(n = 1, 2, \ldots)$ равностепенно непрерывна на отрезке $[x_0, x_0 + \alpha]$. В самом деле, соотношение

(8) дает

$$|\varphi_n(x'') - \varphi_n(x')| \leqslant$$
$$\leqslant M\left\{x'' - x_{l-1} + (l-k-1)\frac{\alpha}{n} + x_k - x'\right\} =$$
$$= M\{x'' - x_{l-1} + (x_{l-1} - x_k) + x_k - x'\} = M(x'' - x'), \tag{9}$$

откуда непосредственно видно, что при достаточно малом $|x'' - x'|$ мы имеем $|\varphi_n(x'') - \varphi_n(x')| < \varepsilon$, где бы ни были расположены точки x' и x'' и каково бы ни было n.

3) Так как последовательность функций $\varphi_n(x)$ ограничена и равностепенно непрерывна на отрезке $[x_0, x_0 + \alpha]$, то в силу доказанной леммы она должна содержать равномерно сходящуюся на этом отрезке подпоследовательность; так как в дальнейшем мы будем иметь дело только с этой подпоследовательностью, то мы можем без всяких опасений обозначить ее по-прежнему через $\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x), \ldots$; таким образом, $\varphi_n(x) \to \varphi(x)$ при $n \to \infty$ равномерно для $x_0 \leqslant x \leqslant x_0 + \alpha$. Покажем теперь, что определенная таким образом функция $\varphi(x)$ удовлетворяет всем требованиям доказываемой теоремы. Так как $\varphi_n(x_0) = y_0$ при любом n, то, прежде всего, $\varphi(x_0) = y_0$.

4) Убедимся, что функция $\varphi(x)$ при $x_0 \leqslant x \leqslant x_0 + \alpha$ удовлетворяет дифференциальному уравнению (3). Пусть задано произвольное $\varepsilon > 0$. В силу непрерывности функции $f(x, y)$ и ограниченности и замкнутости области D существует такое $\delta > 0$, что для любых точек (ξ_1, η_1) и (ξ_2, η_2) области D, связанных неравенствами $|\xi_2 - \xi_1| < \delta$, $|\eta_2 - \eta_1| < 2M\delta$, выполняется и неравенство

$$|f(\xi_2, \eta_2) - f(\xi_1, \eta_1)| < \varepsilon.$$

В силу неравенства (9) мы имеем при $x' \leqslant x_i \leqslant x''$, полагая, как прежде, $\varphi_n(x_i) = y_i$,

$$|y_i - \varphi_n(x')| \leqslant M(x_i - x') \leqslant M(x'' - x');$$

а так как при достаточно большом n

$$|\varphi_n(x') - \varphi(x')| \leqslant M(x'' - x'),$$

то
$$|y_i - \varphi(x')| \leqslant 2M(x'' - x');$$
поэтому, если $|x'' - x'| \leqslant \delta$, то при достаточно большом n и $x' \leqslant x_i \leqslant x''$ мы имеем
$$|x_i - x'| < \delta, \ |y_i - y'| \leqslant 2M\delta,$$
где положено $y' = \varphi(x')$. Отсюда по определению числа δ мы получаем при $x' \leqslant x_i \leqslant x''$ и достаточно большом n
$$|f(x_i, y_i) - f(x', y')| < \varepsilon,$$
или
$$f(x', y') - \varepsilon \leqslant f(x_i, y_i) \leqslant f(x', y') + \varepsilon;$$
если мы применим эту оценку к каждому члену правой части формулы (8), то мы, очевидно, получим, что при $0 < x'' - x' < \delta$ и достаточно большом n
$$[f(x', y') - \varepsilon](x'' - x') \leqslant$$
$$\leqslant \varphi_n(x'') - \varphi_n(x') \leqslant [f(x', y') + \varepsilon](x'' - x'),$$
а так как левая и правая части этих неравенств от n не зависят, что в пределе мы получаем, что при единственном условии $|x'' - x'| < \delta$
$$[f(x', y') - \varepsilon](x'' - x') \leqslant \varphi(x'') - \varphi(x') \leqslant$$
$$\leqslant [f(x', y') + \varepsilon](x'' - x'),$$
или, что то же,
$$\left|\frac{\varphi(x'') - \varphi(x')}{x'' - x'} - f(x', y')\right| \leqslant \varepsilon.$$

Но это и значит, что
$$\varphi'(x) = f(x, y) \qquad (x_0 \leqslant x \leqslant x_0 + \alpha)$$
(причем $\varphi'(x)$ при $x = x_0$ означает, конечно, правую производную).

5) Так как, наконец, мы, очевидно, можем совершенно аналогично получить тот же результат для отрезка $[x_0 - \alpha, x_0]$, то $\varphi(x)$ удовлетворяет всем требованиям теоремы, которая, таким образом, доказана.

Определенные нами функции $y = \varphi_n(x)$ имеют своими графиками ломаные линии; направление каждого звена такой ломаной совпадает с направлением поля в

его левом конце; длины звеньев с ростом n безгранично уменьшаются; таким образом, чем больше n, тем в более густой сети точек направление ломаной совпадает с направлением поля; мы поэтому совершенно естественно ожидаем, что если эти ломаные имеют предельную кривую (при $n \to \infty$), то направление этой кривой будет совпадать с направлением поля уже в *каждой* ее точке, т. е. это будет интегральная кривая. Существование предельной функции (если не для всей последовательности функций $\varphi_n(x)$, то по меньшей мере для некоторой ее подпоследовательности, что для нашей цели безразлично) устанавливается в предварительной лемме. В дальнейшем нам оставалось только строгим и точным рассуждением подтвердить правильность нашего ожидания, что нами и было сделано.

Единственность решения. Важность этой теоремы, гарантирующей при известных условиях существование решения для уравнения (3), сама собой очевидна. Когда перед нами практически встает необходимость решить какое-нибудь дифференциальное уравнение, мы тратим на это некоторые, иногда весьма значительные, усилия: мы пытаемся свести его решение к «квадратурам» (простым интеграциям), а если это не удается, то применяем обычно те или другие приемы приближенного вычисления; однако, для того чтобы эта наша активность стала разумной, надо, конечно, иметь уверенность, что искомый нами предмет действительно существует; наши старания являли бы собой жалкую картину, если бы наше уравнение вовсе не имело решений.

Не менее важен, конечно, и вопрос о *единственности* искомого решения. Если, положим, поставленная нами задача привела к заключению, что искомая функция $y = \varphi(x)$ должна удовлетворять уравнению (3) и обращаться в y_0 при $x = x_0$, и если мы такое решение уравнения (3) нашли, то мы можем считать нашу задачу решенной только в том случае, если мы уверены, что не существует никакого другого решения уравнения (3), удовлетворяющего тому же «начальному» условию $\varphi(x_0) = y_0$; ибо если бы таких решений существовало несколько, то мы не могли бы иметь уверенности, что нашей задаче отвечает как раз то, которое мы нашли;

и если бы даже мы нашли *все* такие решения уравнения (3), мы все же, вообще говоря, не имели бы оснований для суждения о том, какое из них отвечает на поставленный вопрос.

Весьма замечательно, что при тех условиях, выполнение которых позволило нам выше доказать существование искомого решения (т. е. при постулировании одной только непрерывности функции $f(x, y)$ в данной области), мы не можем еще гарантировать единственности этого решения; бывают случаи, когда функция $f(x, y)$ непрерывна в области D и тем не менее существует несколько решений уравнения (3), обращающихся в y_0 при $x = x_0$. Можно, однако, добиться единственности решения, требуя от функции $f(x, y)$ несколько более того, что дает простая непрерывность; одна из наиболее удобных форм такого усиленного требования гласит: *существует такое постоянное положительное число k, что*

$$|f(x, y_1) - f(x, y_2)| \leqslant k|y_1 - y_2| \tag{A}$$

для любых точек (x, y_1) и (x, y_2) области D[1]; это требование, не будучи, правда, самым широким из возможных, выполняется в большинстве практически встречающихся случаев и по своей простоте очень удобно для применений.

Итак, нужно доказать, что *если функция $f(x, y)$ в области D непрерывна и удовлетворяет условию* (A), *то решение $y = \varphi(x)$ уравнения* (3), *обращающееся в y_0 при $x = x_0$, является единственным.*

Допустим, что каждая из двух функций $\varphi_1(x)$ и $\varphi_2(x)$ удовлетворяет уравнению (3) при $x_0 - \alpha \leqslant x \leqslant x_0 + \alpha$ и что $\varphi_1(x_0) = \varphi_2(x_0) = y_0$. Положим $\varphi_2(x) - \varphi_1(x) = \omega(x)$, так что $\omega(x_0) = 0$. Тогда при $x_0 - \alpha \leqslant x \leqslant x_0 + \alpha$

$$\left|\frac{d\omega}{dx}\right| = \left|\frac{d\varphi_2}{dx} - \frac{d\varphi_1}{dx}\right| = |f(x, \varphi_2(x)) - f(x, \varphi_1(x))| \leqslant \leqslant k\,|\varphi_2(x) - \varphi_1(x)| = k\,|\omega(x)|. \tag{10}$$

---

[1]) Это условие обычно называют условием Липшица.— *Прим. ред.*

Обозначим через μ наибольшее значение функции $|\omega(x)|$ на том из двух отрезков $\left[x_0-\frac{1}{2k}, x_0+\frac{1}{2k}\right]$ и $[x_0-\alpha, x_0+\alpha]$, который меньше другого (пусть это будет отрезок $[x_0-r, x_0+r]$, где, следовательно, r означает меньшее из чисел α и $\frac{1}{2k}$). В силу непрерывности функции $|\omega(x)|$ такое наибольшее значение существует при некотором определенном $x=x_1$: $|\omega(x_1)|=$ $=\mu$. Применяя соотношение (10) и первую теорему о среднем значении, мы находим

$$\mu=|\omega(x_1)|=|\omega(x_1)-\omega(x_0)|=$$
$$=\left|\int_{x_0}^{x_1}\frac{d\omega}{dx}\,dx\right|\leqslant\left|\int_{x_0}^{x_1}\left|\frac{d\omega}{dx}\right|dx\right|^{1)}\leqslant k\left|\int_{x_0}^{x_1}|\omega(x)|\,dx\right|\leqslant$$
$$\leqslant k\mu\cdot|x_1-x_0|\leqslant k\mu\cdot\frac{1}{2k}=\frac{\mu}{2},$$

откуда, конечно, $\mu=0$. Это значит, что $\omega(x)=0$ при $x_0-r\leqslant x\leqslant x_0+r$. Если $r=\alpha$, теорема доказана; если же $r=\frac{1}{2k}<\alpha$, то во всяком случае $\omega\left(x_0-\frac{1}{2k}\right)=$ $=\omega\left(x_0+\frac{1}{2k}\right)=0$; мы можем поэтому повторить наше рассуждение, принимая вместо $x=x_0$ за исходную точку последовательно точки $x_0-\frac{1}{2k}$ и $x_0+\frac{1}{2k}$; это позволит утверждать, что $\omega(x)=0$ уже на отрезке $\left[x_0-2\cdot\frac{1}{2k}, x_0+2\cdot\frac{1}{2k}\right]$ или $[x_0-\alpha, x_0+\alpha]$; продолжая этот процесс достаточное число раз, мы, очевидно, расширим отрезок, в котором $\omega(x)=0$, до пределов $[x_0-$ $-\alpha, x_0+\alpha]$, чем теорема и будет доказана.

Зависимость решений от параметров. Если в основании дифференциального уравнения лежит какая-либо конкретная задача, то это уравнение всегда содержит известное число параметров, представляющих значения тех постоянных величин, которые определяют специфи-

---

[1]) Мы берем абсолютное значение интеграла, учитывая возможность $x_1<x_0$.

ческие условия изучаемого явления. Так, в задаче, приведенной в качестве примера в начале этой лекции, такими величинами служат: вместимость сосудов a, скорость истечения жидкости b и первоначальный запас соли в первом сосуде c; все эти три величины, естественно, входят в составленное нами уравнение; точно так же, конечно, и всякое решение этого уравнения будет зависеть от всех этих параметров. Таким образом, желая подчеркнуть эту зависимость, мы должны писать уравнение (3) в виде

$$\frac{dy}{dx} = f(x, y, p_1, p_2, \ldots, p_r)$$

и решение его в виде

$$y = \varphi(x, p_1, p_2, \ldots, p_r),$$

где $p_1, p_2, \ldots, p_r$ — параметры, о которых мы только что говорили.

Нетрудно понять, что для прикладных задач характер зависимости решений полученного дифференциального уравнения от такого рода параметров имеет весьма существенное значение и что в особенности важно установить *непрерывность* этой зависимости. В самом деле, значения параметров обычно получаются как результаты тех или других измерений и потому, как правило, задаются нам не с абсолютной точностью, а лишь приближенно, с некоторой, хотя бы и малой, погрешностью. Поэтому, если бы при ничтожных изменениях параметров значения функции $\varphi(x, p_1, p_2, \ldots, p_r)$ могли значительно меняться, то такого рода решения задачи были бы практически совершенно бесполезными; для практических целей пригодны лишь такие решения, которые при малых изменениях параметров сами претерпевают лишь малые изменения, так что, приближенно зная значения параметров, мы можем приближенно найти и значения этих решений; но точным математическим выражением этого свойства, очевидно, и является непрерывность функций $\varphi(x, p_1, p_2, \ldots, p_r)$ относительно параметров $p_1, p_2, \ldots, p_r$.

Мы покажем теперь, что если функция $f(x, y, p_1, p_2, \ldots, p_r)$ непрерывна относительно какого-либо из

параметров p_i и притом равномерно относительно положения точки (x, y) в области D, то при соблюдении тех условий, в которых мы доказали выше существование и единственность решений уравнения (3), эти решения также будут непрерывными функциями параметра p_i; так как при этом речь идет о каждом параметре в отдельности, то мы нисколько не ограничим общности доказательства, если для сокращения записей допустим, что функция f (а значит, и функция φ) зависит только от одного параметра p.

Итак, допустим, что функция $f(x, y, p)$ непрерывна относительно всех трех своих аргументов, когда точка (x, y) находится в замкнутой ограниченной области D, а параметр p принадлежит некоторому отрезку d, и что, кроме того,

$$|f(x, y_1, p) - f(x, y_2, p)| \leqslant k|y_1 - y_2|, \qquad (A')$$

если $(x, y_1) \in D$, $(x, y_2) \in D$ и $p \in d$. Мы утверждаем, что в таком случае то единственное решение $y = \varphi(x, p)$ уравнения

$$\frac{dy}{dx} = f(x, y, p), \qquad (3'')$$

которое обращается в y_0 при $x = x_0$, непрерывно относительно p на отрезке d.

Для доказательства нам придется вернуться к той конструкции, с помощью которой мы выше доказали существование решения уравнения (3), и исследовать характер зависимости от параметра p построенных нами там функций $\varphi_n(x)$, которые мы теперь будем, конечно, писать в виде $\varphi_n(x, p)$. Величины y_i, определяемые с помощью рекуррентного соотношения (5), также, очевидно, зависят от p, $y_i = y_i(p)$, и формула (5) принимает вид

$$y_i(p) = y_{i-1}(p) + f(x_{i-1}, y_{i-1}(p), p)\frac{a}{n}. \qquad (11)$$

В силу предположенной нами непрерывности функции $f(x, y, p)$ относительно параметра p для любого $\varepsilon > 0$ существует такое $\delta > 0$, что

$$|f(x, y, p + h) - f(x, y, p)| < \varepsilon \qquad (12)$$

при $|h| < \delta$, $p \in d$, $p + h \in d$, $(x, y) \in D$. В силу формулы (11)

$$y_i(p+h) - y_i(p) = y_{i-1}(p+h) - y_{i-1}(p) + \\ + \frac{\alpha}{n}\{f(x_{i-1}, y_{i-1}(p+h), p+h), - \\ - f(x_{i-1}, y_{i-1}(p), p)\} \\ (1 \leqslant i \leqslant n).$$

Положим для краткости

$$y_i(p+h) - y_i(p) = \Delta_i \quad (0 \leqslant i \leqslant n),$$

так что

$$\Delta_i = \Delta_{i-1} + \frac{\alpha}{n}\{f(x_{i-1}, y_{i-1}(p+h), p+h) - \\ - f(x_{i-1}, y_{i-1}(p), p)\} \quad (1 \leqslant i \leqslant n). \tag{13}$$

Выражение, стоящее в фигурных скобках, можно представить в виде

$$f(x_{i-1}, y_{i-1}(p+h), p+h) - f(x_{i-1}, y_{i-1}(p), p+h) + \\ + f(x_{i-1}, y_{i-1}(p), p+h) - f(x_{i-1}, y_{i-1}(p), p).$$

В силу (A′) первая из этих разностей по абсолютному значению не превосходит $k|\Delta_{i-1}|$, а в силу (12) вторая разность по абсолютному значению меньше чем ε; таким образом, мы получаем

$$|\Delta_i| < |\Delta_{i-1}| + \frac{\alpha}{n}\{k|\Delta_{i-1}| + \varepsilon\} = \frac{\varepsilon\alpha}{n} + |\Delta_{i-1}|\left(1 + \frac{\alpha k}{n}\right)$$

$(1 \leqslant i \leqslant n)$. Применяя эту же оценку к стоящей в правой части величине $|\Delta_{i-1}|$ и повторяя этот процесс, мы придем к соотношению

$$|\Delta_i| < \frac{\varepsilon\alpha}{n} + \frac{\varepsilon\alpha}{n}\left(1 + \frac{\alpha k}{n}\right) + \frac{\varepsilon\alpha}{n}\left(1 + \frac{\alpha k}{n}\right)^2 + \ldots \\ \ldots + \frac{\varepsilon\alpha}{n}\left(1 + \frac{\alpha k}{n}\right)^{i-1} = \frac{\varepsilon}{k}\left[\left(1 + \frac{\alpha k}{n}\right)^i - 1\right] \leqslant \\ \leqslant \frac{\varepsilon}{k}\left[\left(1 + \frac{\alpha k}{n}\right)^n - 1\right] < \frac{\varepsilon}{k}\left(e^{\alpha k} - 1\right) \quad (1 \leqslant i \leqslant n).$$

Таким образом, при $|h| < \delta$ мы имеем

$$|y_i(p+h) - y_i(p)| < \frac{\varepsilon}{k}\left(e^{\alpha k} - 1\right) \quad (1 \leqslant i \leqslant n),$$

или, что то же,

$$|\varphi_n(x_i, p+h) - \varphi_n(x_i, p)| < \frac{\varepsilon}{k}\left(e^{\alpha k} - 1\right)$$

$$(n = 1, 2, \ldots;\ 1 \leqslant i \leqslant n);$$

но на каждом отрезке $[x_{i-1}, x_i]$ функции $\varphi_n(x, p)$ и $\varphi_n(x, p+h)$ линейны; поэтому неравенство

$$|\varphi_n(x, p+h) - \varphi_n(x, p)| < \frac{\varepsilon}{k}\left(e^{\alpha k} - 1\right),$$

выполняющееся, как мы только что видели, при $|h| < \delta$ в концах каждого такого отрезка, по необходимости должно выполняться и на всем этом отрезке, а значит, на всем отрезке $[x_0 - \alpha, x_0 + \alpha]$. Ввиду произвольности числа ε это означает, что последовательность функций $\varphi_n(x, p)$ равностепенно непрерывна относительно параметра p на отрезке d. Обозначим через S эту последовательность; как мы видели прежде, последовательность S содержит подпоследовательность S', сходящуюся равномерно относительно x на отрезке $[x_0 - \alpha, x_0 + \alpha]$ к функции $\varphi(x, p)$, которая и является искомым решением уравнения (3″). Но последовательность S', будучи частью последовательности S, очевидно, также есть последовательность, равностепенно непрерывная относительно параметра; поэтому в силу прежде доказанной леммы она должна содержать подпоследовательность S'', сходящуюся (конечно, к той же функции $\varphi_n(x, p)$) равномерно относительно p на отрезке d; а так как все функции $\varphi_n(x, p)$ непрерывны относительно p, то отсюда следует, что и функция $\varphi(x, p)$ непрерывна относительно p на отрезке d, что и требовалось доказать.

Как мы знаем, каждое решение $\varphi(x)$ уравнения (3) в наших условиях однозначно определяется тем значением y_0, которое оно принимает при $x = x_0$; функция $\varphi(x)$ меняется при изменении чисел x_0 и y_0, так что, по сути дела, она является функцией $\varphi(x, x_0, y_0)$ трех

независимых переменных; по тем же причинам, о которых мы говорили выше, характер зависимости решения $\varphi(x, x_0, y_0)$ от этих «начальных данных» x_0 и y_0 имеет существенное значение, и притом не только теоретическое, но и практическое. На первый взгляд может показаться, что это — новая задача, которая не сводится к только что рассмотренному нами вопросу о зависимости решений от параметров, ибо числа x_0 и y_0 явно не входят в качестве аргументов в функцию $f(x, y)$. На самом деле, однако, такая редукция вполне возможна. В самом деле, если мы преобразуем в уравнении (3) как независимую переменную, так и искомую функцию с помощью соотношений

$$x = x_0 + x^*, \quad y = y_0 + y^*,$$

где x^* и y^* — новые независимая переменная и искомая функция, то уравнение (3) получит вид

$$\frac{dy^*}{dx^*} = f(x_0 + x^*, y_0 + y^*), \qquad (3''')$$

причем требуется найти решение этого уравнения, обращающееся в нуль при $x^* = 0$. Здесь уже x_0 и y_0 явно фигурируют в качестве параметров в правой части уравнения; пусть

$$y^* = \varphi^*(x^*, x_0, y_0) \qquad (14)$$

есть искомое решение уравнения $(3''')$. Оно непрерывно относительно x_0 и y_0 в силу только что доказанной нами теоремы, так как функция $f(x_0 + x^*, y_0 + y^*)$, будучи непрерывной относительно каждого из своих двух аргументов, тем самым автоматически непрерывна относительно x_0 и y_0. Но чтобы получить искомое решение уравнения (3), мы, очевидно, должны только перейти в выражении (14) от новых переменных к старым, что дает

$$y = \varphi(x, x_0, y_0) = y_0 + \varphi^*(x - x_0, x_0, y_0),$$

откуда непосредственно вытекает, что это решение непрерывно зависит от начальных значений x_0 и y_0.

Преобразование переменных. Вы знаете, что при интегрировании функций одним из самых мощных приемов упрощения задачи, иногда приводящим и к полно-

му ее разрешению, служит преобразование переменной интеграции (так называемый «метод подстановки»). При решении дифференциальных уравнений этот метод также является одним из наиболее сильных орудий; гибкость его здесь еще значительно увеличивается тем обстоятельством, что преобразованию может быть подвергнута как независимая переменная, так и искомая функция. Как мы видели в начале нашей лекции, решение дифференциального уравнения легко приводится к квадратурам, если переменные могут быть «разделены» (это означает, что посредством умножений и делений уравнение приводится к виду $M(y)\,dy+N(x)\,dx=0$, где в каждом из двух слагаемых левой части фигурирует только одна из переменных x, y); именно этой цели и служит обычно преобразование переменных; преобразуя либо независимую переменную x, либо искомую функцию y, либо ту и другую одновременно, нам часто удается заменить уравнение, в котором переменные не разделялись, новым, в котором они уже могут быть разделены. Хотя класс дифференциальных уравнений, которые этим путем могут быть сведены к квадратурам, и очень ограничен, все же он содержит целый ряд простейших, а потому и наиболее часто встречающихся на практике типов, вследствие чего метод преобразования переменных получает большое практическое значение.

Сюда относятся прежде всего, среди уравнений первого порядка, все *линейные* уравнения, т. е. уравнения, в которых как неизвестная функция y, так и ее производная y' фигурируют лишь в первой степени; общий вид такого уравнения:

$$y' + f_1(x)\,y + f_2(x) = 0, \qquad (15)$$

где $f_1(x)$ и $f_2(x)$ — данные непрерывные на некотором отрезке (a, b) функции от x. Все уравнения этого типа могут быть с помощью одного и того же приема приведены к виду, в котором переменные допускают разделение.

Чтобы в этом убедиться, рассмотрим сначала уравнение

$$z' + f_1(x)\,z = 0, \qquad (16)$$

где через z обозначена искомая функция; это уравнение — того же типа, что уравнение (15), но только без «свободного члена» $f_2(x)$; переменные в нем немедленно разделяются:

$$\frac{dz}{z} + f_1(x)\,dx = 0,$$

и интеграция дает

$$\ln\frac{z}{z_0} + \int_{x_0}^{x} f_1(u)\,du = 0,$$

где x_0 — какая-нибудь точка отрезка $(a,\ b)$; для наших целей достаточно иметь какое-нибудь одно решение уравнения (16), поэтому мы положим $z_0 = 1$, так что

$$\ln z + \int_{x_0}^{x} f_1(u)\,du = 0,$$

откуда

$$z = e^{-\int_{x_0}^{x} f_1(u)du} = \varphi(x); \tag{17}$$

таким образом, решение уравнения (16) находится с помощью одной квадратуры.

Чтобы свести к квадратурам и общее уравнение (15), преобразуем в нем искомую функцию y с помощью соотношения

$$y = \varphi(x)\,y^*,$$

где y^* — новая искомая функция, а $\varphi(x)$ — определяемое формулой (17) решение уравнения (16); тогда уравнение (15) принимает вид

$$\varphi(x)\,(y^*)' + \varphi'(x)\,y^* + f_1(x)\,\varphi(x)\,y^* + f_2(x) = 0,$$

или

$$\varphi(x)\,(y^*)' + f_2(x) = 0,$$

так как в силу уравнения (16)

$$\varphi'(x) + f_1(x)\varphi(x) = 0;$$

отсюда

$$(y^*)' = -\frac{f_2(x)}{\varphi(x)} = -f_2(x)\, e^{\int\limits_{x_0}^{x} f_1(u)du},$$

и, следовательно,

$$y^* = -\int\limits_{x_0}^{x} f_2(v)\, e^{\int\limits_{x_0}^{v} f_1(u)du}\, dv + C,$$

где C — постоянная интеграции; наконец,

$$y = \varphi(x)\, y^* = -e^{-\int\limits_{x_0}^{x} f_1(u)du} \left\{ \int\limits_{x_0}^{x} f_2(v)\, e^{\int\limits_{x_0}^{v} f_1(u)du}\, dv + C \right\}.$$

Это — общее решение уравнения (15), из которого надлежащим выбором постоянной C можно получить все его частные решения; если мы хотим, например, иметь $y = y_0$ при $x = x_0$, то из общего решения находим $y_0 = -C$, и искомое частное решение получает вид

$$y = e^{-\int\limits_{x_0}^{x} f_1(u)du} \left\{ y_0 - \int\limits_{x_0}^{x} f_2(v)\, e^{\int\limits_{x_0}^{v} f_1(u)du}\, dv \right\}. \qquad (18)$$

Мы видим, таким образом, что решение уравнения (15) с помощью надлежаще выбранного преобразования искомой функции сводится к двум последовательным квадратурам. В качестве примера доведем до конца решение задачи, рассмотренной в начале этой лекции.

Мы пришли там к уравнению

$$\frac{dy}{dt} + \frac{b}{a}y - \frac{bc}{a}e^{-\frac{b}{a}t} = 0,$$

которое, очевидно, есть уравнение типа (15); здесь $f_1(t) = \frac{b}{a}$ и $f_2(t) = -\frac{bc}{a}e^{-\frac{b}{a}t}$; кроме того, $t_0 = 0$ и $y_0 = 0$, так как в момент $t = 0$ второй сосуд вовсе не

содержит соли. Мы имеем $\int\limits_0^t f_1(u)\,du = \frac{b}{a}t$, и формула (18) дает

$$y = e^{-\frac{b}{a}t}\left\{0 + \int\limits_0^t \frac{bc}{a} e^{-\frac{b}{a}v} e^{\frac{b}{a}v}\,dv\right\} = \frac{bc}{a} te^{-\frac{b}{a}t}. \quad (19)$$

Эта формула полностью решает поставленную задачу. Исследуя найденную функцию, мы легко находим, что количество соли во втором сосуде сначала возрастает, а затем с момента $t=\frac{a}{b}$ начинает убывать и стремится к нулю при $t\to+\infty$; в момент $t = \frac{a}{b}$, когда количество соли во втором сосуде наибольшее, это количество равно $\frac{c}{e}$; замечательно, что это максимальное количество не зависит ни от a, ни от b; наоборот, промежуток времени, необходимый для максимального насыщения солью второго сосуда, зависит, как мы видим, от a и b, но не зависит от c. Все эти и многие другие интегральные черты рассматриваемого явления дает нам изучение функции (19); мы видим, что решение дифференциального уравнения в данном случае действительно позволяет нам получить все необходимые сведения о течении процесса в целом, в то время как само дифференциальное уравнение дает нам лишь *мгновенные* (*локальные*) соотношения между участвующими в этом процессе величинами.

Другой часто встречающийся тип уравнений, простым преобразованием неизвестной функции приводящихся к уравнениям с разделяющимися переменными, представляют собой так называемые *однородные* уравнения, общий вид которых

$$\frac{dy}{dx} = f\left(\frac{y}{x}\right); \quad (20)$$

в частности, к этому типу принадлежат часто встречающиеся уравнения вида

$$P(x, y)\,dy = Q(x, y)\,dx,$$

где $P(x, y)$ и $Q(x, y)$ — однородные многочлены одной и той же степени n. В самом деле, такой однородный многочлен, как легко видеть, после деления на x^n становится многочленом относительно переменной $\frac{y}{x}$, вследствие чего в правой части соотношения

$$\frac{dy}{dx} = \frac{Q(x, y)}{P(x, y)} = \frac{P(x, y) : x^n}{Q(x, y) : x^n}$$

числитель и знаменатель — целые рациональные функции отношения $\frac{y}{x}$, а значит, и вся правая часть — рациональная функция этого отношения.

Преобразуя в уравнении (20) неизвестную функцию с помощью соотношения $y = xy^*$, мы приводим его к виду

$$y^* + x(y^*)' = f(y^*),$$

откуда $\frac{dy^*}{dx} = \frac{f(y^*) - y^*}{x}$, или

$$\frac{dy^*}{f(y^*) - y^*} = \frac{dx}{x};$$

переменные разделены; интегрируя, легко находим выражение x через y^*; если это соотношение однозначно разрешимо относительно y^*, то из него мы получаем, обратно, выражение y^*, а значит, и y через x; в общем случае это соотношение определяет y^*, а следовательно, и y, как «неявную» функцию от x.

Системы уравнений и уравнения высших порядков. В заключение коснемся еще кратко вопросов о системах дифференциальных уравнений первого порядка и об уравнениях высших порядков.

Пусть мы имеем систему n дифференциальных уравнений первого порядка, содержащую n неизвестных функций $y_1, y_2, \ldots, y_n$ от одной независимой переменной x; само собой разумеется, что *решением* такой системы мы назовем любую систему функций $y_i = \varphi_i(x)$ $(1 \leqslant i \leqslant n)$, удовлетворяющих всем данным уравнениям. Мы допустим, что данная система уравнений разре-

шена относительно производных $\frac{dy_i}{dx}$ $(1 \leqslant i \leqslant n)$ и, следовательно, имеет вид

$$\frac{dy_i}{dx} = f_i(x, y_1, y_2, \ldots, y_n) \qquad (1 \leqslant i \leqslant n). \qquad (21)$$

Совокупность значений переменных x, y_1, …, y_n нам удобно и здесь называть «точкой»; это, конечно, точка в пространстве $n+1$ измерений; мы будем во всем дальнейшем предполагать все функции f_i непрерывными в некоторой области D этого пространства. Всякая система функций $y_i = \varphi_i(x)$ $(1 \leqslant i \leqslant n)$ имеет своим геометрическим образом в этом пространстве некоторую кривую; если данные функции образуют собой решение системы (21), то изображающую их кривую мы будем называть интегральной кривой системы (21). Эта геометрическая терминология чрезвычайно полезна здесь не только своей наглядностью, но и, главным образом, тем, что с ее помощью многие формулировки и рассуждения могут быть проведены и изложены для системы уравнений в тех же схемах и терминах, как и в случае одного уравнения.

Так, основную теорему существования решений мы можем формулировать совсем кратко, говоря, что через каждую точку области D проходит по меньшей мере одна интегральная кривая; аналитически это означает, конечно, что, какова бы ни была система $n+1$ чисел $x_0, y_1^{(0)}, y_2^{(0)}, y_n^{(0)}$, принадлежащая области D, существует система функций $y_i = \varphi_i(x)$ $(1 \leqslant i \leqslant n)$, удовлетворяющих системе уравнений (21) на некотором отрезке $x_0 - \alpha \leqslant x \leqslant x_0 + \alpha$, причем $\varphi_i(x_0) = y_i^{(0)}$ $(1 \leqslant i \leqslant n)$. Доказательство этой теоремы для системы уравнений только в чисто техническом отношении сложнее, чем для случая одного уравнения. В основании его лежит лемма, в точности аналогичная той, которой мы пользовались прежде; эта лемма может быть либо доказана непосредственно, либо выведена как следствие нашей прежней леммы. Последний способ особенно прост; элементом последовательности здесь является не отдельная функция, а система из n функций $\{F_1(x), F_2(x), \ldots \ldots, F_n(x)\} = S$, и надо доказать, что в предположении

ограниченности и равностепенной непрерывности множества всех функций, входящих во все системы данной последовательности, мы можем выбрать последовательность систем $S_1, S_2, \ldots, S_k, \ldots$, равномерно сходящуюся на данном отрезке. Это означает, что, полагая

$$S_k = \{F_{1,k}(x),\ F_{2,k}(x),\ \ldots,\ F_{n,k}(x)\} \quad (k = 1, 2, \ldots),$$

мы получаем n функциональных последовательностей

$$F_{i,k}(x) \quad (k = 1, 2, \ldots) \quad (1 \leqslant i \leqslant n),$$

каждая из которых сходится равномерно на данном отрезке. Доказательство строится так: сначала в силу нашей прежней леммы мы находим такую последовательность систем S_k, чтобы последовательность $F_{1,k}(x)$ равномерно сходилась на данном отрезке; из этой последовательности систем мы в силу той же леммы можем выбрать такую подпоследовательность, в которой равномерно сходилась бы на данном отрезке и последовательность вторых функций $F_{2,k}(x)$; повторяя этот процесс n раз, мы, очевидно, приходим к такой последовательности систем, в которой уже все n последовательностей $F_{i,k}(x)$ $(1 \leqslant i \leqslant n)$ равномерно сходятся на данном отрезке, чем и завершается доказательство новой, расширенной леммы.

Дальнейшее доказательство основной теоремы существования проводится в тесной аналогии с нашим прежним рассуждением; в качестве руководящей нити здесь удобнее всего принять нашу геометрическую иллюстрацию; здесь, как и прежде, мы строим некоторое множество ломаных линий со сколь угодно малыми длинами звеньев и из этого множества в силу доказанной леммы выбираем последовательность, равномерно сходящуюся к некоторой кривой, относительно которой может быть доказано прежним методом (с небольшими и самоочевидными усложнениями чисто технического характера), что она есть интегральная кривая системы (21); эта кривая проходит через данную точку $(x_0, y_1^{(0)}, \ldots, y_n^{(0)})$, так как через нее проходят все построенные ломаные.

Единственность проходящей через данную точку интегральной кривой может быть, как и прежде, доказана лишь при наложении на функции f_i некоторых дополнительных условий; простейшая форма этих условий вполне аналогична условию (A) и сводится к требованию, чтобы во всей области D выполнялись неравенства

$$\left|f_i\left(x, y_1^{(1)}, y_2^{(1)}, \ldots, y_n^{(1)}\right) - f_i\left(x, y_1^{(2)}, y_2^{(2)}, \ldots, y_n^{(2)}\right)\right| \leqslant$$
$$\leqslant k \sum_{i=1}^{n} \left|y_i^{(2)} - y_i^{(1)}\right| \quad (1 \leqslant k \leqslant n),$$

где k — некоторое постоянное положительное число.

У нас не осталось времени для уравнений порядка выше первого; покажем лишь, что решение всякого уравнения

$$F(x, y, y', y'', \ldots, y^{(n)}) = 0 \qquad (22)$$

порядка n может быть сведено к решению системы уравнений первого порядка. В самом деле, рассмотрим систему n уравнений первого порядка, с неизвестными функциями $y_1, y_2, \ldots, y_n$, следующего вида:

$$\left.\begin{aligned} F\left(x, y_1, y_2, \ldots, y_{n-1}, y_n, y_n'\right) = 0,\\ y_1' = y_2,\\ y_2' = y_3,\\ \ldots\ldots\\ y_{n-1}' = y_n, \end{aligned}\right\} \qquad (23)$$

и допустим, что мы нашли решение

$$y_i = \varphi_i(x) \qquad (1 \leqslant i \leqslant n)$$

этой системы; в силу уравнений (23) тогда

$$\varphi_1'(x) = \varphi_2(x) = y_2,$$
$$\varphi_1''(x) = \varphi_2'(x) = \varphi_3(x) = y_3,$$
$$\varphi_1'''(x) = \varphi_2''(x) = \varphi_3'(x) = \varphi_4(x) = y_4,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\varphi_1^{(n-1)}(x) = \varphi_2^{(n-2)}(x) = \ldots = \varphi_n(x) = y_n, \ \varphi_1^{(n)}(x) = y_n';$$

поэтому первое из уравнений (23) дает

$$F\left(x, \varphi_1(x), \varphi_1'(x), \ldots, \varphi_1^{(n)}(x)\right) = 0,$$

т. е. функция $y = \varphi_1(x)$ есть решение уравнения (22).

Обратно, пусть функция $y = \varphi(x)$ удовлетворяет уравнению (22). Полагая

$$y_1 = \varphi(x),\ y_2 = \varphi'(x),\ \ldots,\ y_n = \varphi^{(n-1)}(x),$$

мы непосредственно видим, что система функций $(y_1, y_2, \ldots, y_n)$ образует решение системы (23). Таким образом, задачи отыскания совокупности решений уравнения (22) и системы (23) действительно полностью сводятся друг к другу.

Александр Яковлевич Хинчин

ВОСЕМЬ ЛЕКЦИЙ
ПО МАТЕМАТИЧЕСКОМУ АНАЛИЗУ

М., 1977 г., 280 стр. с илл.

Редакторы *М. А. Крейнес, Ф. И. Кизнер,*

Техн. редактор *Л. В. Лихачева*

Корректор *В. П. Сорокина*

Сдано в набор 22/IX 1976 г. Подписано к печати 28/I 1977 г. Бумага $84\times108^{1}/_{32}$. Физ. печ. л. 8,75. Условн. печ. л. 14,7. Уч.-изд. л. 12,95. Тираж 60 000 экз. Т-04109. Цена книги 45 коп. Заказ № 275

Издательство «Наука»
Главная редакция
физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

4-я типография издательства «Наука».
630077, Новосибирск, 77, Станиславского, 25.

Цена 45 коп.